150.8 A673 no.161 c.1
Woodworth, Robert Sessions,
Archives of psychology. --
R.W.B. JACKSON LIBRARY
OISE CIR
3 0005 03037 6266

THE ONTARIO INSTITUTE FOR STUDIES IN EDUCATION LIBRARY

Verbal, Numerical and Spatial Abilities of Young Children

BY
BELLE SCHILLER, Ph.D.

ARCHIVES OF PSYCHOLOGY
R. S. WOODWORTH, EDITOR
No. 161

ONTARIO COLLEGE OF EDUCATION
DEPT. OF [illegible]
UNIVERSITY OF TORONTO

).8
73
161
4

NEW YORK
March, 1934

ARCHIVES OF PSYCHOLOGY

COLUMBIA UNIVERSITY, NEW YORK CITY

The Subscription price is six dollars per volume of about 500 pages. Volume I comprises
Nos. 2-1
Nos. 33-
IX, Nos.
XIII, Nos
Volume X
XX, Nos.
147-154.

5. Fiel
70c.
6. Infl
Acti
7. Stat
Wei
8. Mag
acti
9. Perc
Ham
10. Tim
Bro
11. Hea
Bru
12. Stud
E. A
13. Inac
wor
14. Qua
row.
15. Psyc
$1.2
16. Elec
and
tions
17. Rela
Stro
18. Atte
Time
19. Emp
divid
$1.25
20. Visu
Colo
(Cl.,
21. Curv
(Cl.,
23. Reac
T. Po
24. Inter
Cull
25. React
60c.
26. Study
ers.
27. Stati
S. Ca
28. Ment
M. J.
29. Expe
Holl
30. Psych
Keen
Pupil
31. Fatig
E. As
32. Trans
A. Ma
33. Intell
Public
Thor
(Cl.,
34. Quick
D. O.
35. Overc
Morga
36. Psych
Jr.
37. Effect
J. E.
38. Effect
Gener
(Cl.,
39. Mecha
M. A.
41. Menta
ercise
85c.
42. Psych
fringement: R. H. Paynter. 85c. (Cl., $1.10.)

tions: C. B. Key. $1.25.
84. Consistency of Rate of Work: C. E. Dowd. $1.00.

(*Continued on inside back cover.*)

THE LIBRARY

The Ontario Institute
for Studies in Education

Toronto, Canada

Verbal, Numerical and Spatial Abilities of Young Children

BY
BELLE SCHILLER, Ph.D.

LIBRARY
JUN 3 1969
THE ONTARIO INSTITUTE
FOR STUDIES IN EDUCATION

ARCHIVES OF PSYCHOLOGY
R. S. WOODWORTH, EDITOR
No. 161

ONTARIO COLLEGE OF EDUCATION
DEPT. OF EDUCATIONAL RESEARCH
UNIVERSITY OF TORONTO

NEW YORK
March, 1934

TABLE OF CONTENTS

CHAPTER PAGE

I. THE PROBLEM AND ITS SETTING 5
  1. Introduction 5
  2. Problem 5
  3. Limitations of Previous Research 6
  4. Historical Survey 6
  5. Summary 11

II. TESTING SET-UP 12
  1. Problem 12
  2. Tests 12
  3. Subjects 16
  4. Testing Procedure 18
    A. Order of Tests 18
    B. Administration of Tests 18
    C. Motivation and Rapport 19

III. PRELIMINARY STATISTICAL RESULTS 20
  1. Scoring 20
  2. Reliability Coefficients 20
  3. Distributions 22
  4. Sex Differences 23
    A. Ability 23
    B. Variability 24
  5. Differential Bilingualism 25

IV. ANALYSIS OF CORRELATIONAL RESULTS 27
  1. Preliminary Treatment of Correlations 27
  2. Probable Errors 31
  3. Average Intercorrelations 31

V. FACTOR ANALYSIS OF CORRELATIONAL RESULTS 36
  1. Tetrads within each Test-Category 36
    A. The Verbal Tests 36
    B. The Number Tests 37
    C. The Spatial Tests 38
  2. Central Factor Correlations 40
  3. Cross-Category Tetrads 42
    A. Verbal-Numerical 42
    B. Verbal-Spatial 43
    C. Numerical-Spatial 44
  4. Mean Tetrads 45
    A. Verbal-Numerical 46
    B. Verbal-Spatial 52
    C. Numerical-Spatial 52
  5. The Results of the Several Methods of Tetrad Analysis 53

VI. SUMMARY AND IMPLICATIONS 58

BIBLIOGRAPHY 61

APPENDICES 64
  I. Questionnaire 64
  II. Testing Schedule and Examiners 65
  III. Frequency Distributions 67
  IV. Performance Sub-Test Intercorrelations 69

CHAPTER I[1]

THE PROBLEM AND ITS SETTING

1. INTRODUCTION

The assumption of the existence of general mental ability as a unitary trait, capable of being measured by group omnibus tests of intelligence, has been subjected to analysis in an increasing number of investigations. General intelligence tests are characterized not only by differences in underlying principles of construction and administrative techniques, but by a diversified array of tasks and materials. Nevertheless from these very different agglomerates estimates of general intelligence are derived and are used interchangeably as if they represented equivalent, albeit different, aspects of an identical mental function.

Within each general intelligence test battery there are present sub-tests which may overlap or oppose each other in various ways. If this happens the total score, instead of being an equitable measure of a unitary function, will represent an indiscriminate sampling of abilities some of which may actually be independent and discrete. Thus, depending upon the type of ability emphasized by the test, the individual may be unduly credited or penalized at the expense of his other abilities.

2. PROBLEM

The research to be reported here is concerned with an analysis of representative verbal, numerical, and pictorial or spatial[2] elements constituting group tests designed to measure the general intelligence of young children.

The following problems are investigated:

1. Do the sub-tests within a general intelligence battery measure fundamentally the *same* ability, regardless of the *type* of test material employed?

2. *Within* a given category, *i.e.*, verbal, numerical, spatial, etc., are the test abilities more closely related than they are *between* cate-

---

[1] This is one of a series of studies under the general direction of Prof. H. E. Garrett, who suggested the problem, and for whose unfailing interest and indispensable criticism the writer is deeply grateful.

[2] "Spatial" is used as a general term more closely descriptive of the attributes common to such tests than "pictorial," "concrete," etc.

gories, so that the existence of independent and differentiable mental abilities may be demonstrated?

3. Limitations of Previous Research

The many phases of intelligence testing have been voluminously and often controversially treated. Only a limited amount of data is available, however, in which the necessary experimental and statistical controls have been employed, and a sufficient number of variant abilities tested to enable us to answer the two questions stated above. Differences in tests, subjects and basic assumptions are so great that the resulting contradictions are almost necessarily inevitable.

The following inadequacies which render untrustworthy much of the experimental literature may be mentioned briefly:

1. The N's are small (7, 19, 22, 28, 31, 32, 46, 50, 67, 70).
2. Subjects are heterogeneous with respect to
    a. Age (10, 28, 38, 50)
    b. Sex (10, 19, 28, 31, 33, 38, 46, 50, 58)
    c. School Grade (28, 38, 50)
    d. Socio-economic Conditions (31, 38, 50).
3. Tests are too short or too few for adequate sampling of the various constituent abilities in question (55, 56, 57).
4. The reliabilities are not stated or low (10, 22, 31, 58, 55, 56, 72).
5. Raw r's are averaged when their range is large instead of their transmuted z-values (10, 19, 46).[3]
6. The 5 P.E. criterion as used in the tetrad analysis predisposes the results in favor of the Two-Factor Theory (10, 54).[4]

In the survey which follows reference is limited to research with sizeable groups (*i.e.*, a minimum of at least 100 subjects) and to results obtained for differentiated verbal, number and spatial materials. Data based upon standard intelligence tests in which subtest measures in these abilities were combined into a total score; and data which included abilities other than those involved in intelligence tests are not here discussed.

4. Historical Survey

The characteristic form of much of the present-day treatment of intelligence-test results and mental organization is to a great extent due to the theory and technique of C. Spearman. In 1904 (50) Spearman analyzed the correlational results for the tests and subjects described in Table I by means of three methods which he then introduced—the correlation of common elements, the examination of the correlation-table for hierarchal order, and the calculation of the ratio of the common to the specific factor.[5]

---

[3] *Cf.* Fisher (14).

[4] *Cf.* Peatman (42).

TABLE I

SUBJECTS AND TESTS USED BY SPEARMAN IN 1904 (50)

Subjects	*Tests*
24 village school-children 11 yrs. 2 mos.–13 yrs. 8 mos.	Pitch, light and weight discrimination; rank for common sense, school cleverness
36 children in same school 5 yrs. 6 mos.–10 yrs. 2 mos.	Pitch discrimination, rating for ''intellectual class''
23 boys in ''high class preparatory school'' 9 yrs. 5 mos.–13 yrs. 7 mos.	Pitch, light and weight discrimination; school rank in Classics, French, Math., English
33 boys in ''high class preparatory school'' 9 yrs. 5 mos.–13 yrs. 7 mos.	Pitch discrimination; rank in Classics, French, English, Math., Music
27 adults 21–78 years	Pitch, light and weight discrimination

He concluded that ''All branches of intellectual activity have in common one function (or group of functions) whereas the remaining or specific elements of the activity seem in every case to be wholly different from that in all the others.''

With the extension of the hierarchy-criterion into the more objective method of tetrad-analysis, and on the basis of further data, this conclusion was later elaborated by Spearman into the Two-Factor Theory. This theory postulates a factor ''g'' which enters into all mental operations to a greater or lesser degree; and a factor or factors ''s'' specific to each operation, which varies independently both of ''g'' and of the other ''s'' 's. The factor ''g'' is ''measured (roughly) by current tests of intelligence.''

Overlapping among the ''s'' 's produces group factors. Such overlapping occurs when the abilities differ only in part, or when they are ''very closely and obviously allied.'' Abilities for logical, mechanical, psychological and arithmetical materials, and retentivity comprise the possible group factors conceded by Spearman.

Other unitary factors, which do not involve ''g'' and are themselves general factors of a fundamentally different kind have been postulated. These are w, "a general character factor of purposive consistency and self-control" (69), p, a general factor of "perseveration" or inertia of cognitive processes (36), c, a general factor of cleverness or non-perseveration, "a freedom from inertia" (16), and "o," a general factor of oscillation of the amount of mental output (15).

---

[5] Square of the correlation between the variable and the common factor.

Spearman's early work (27, 50) was challenged in 1913 by Thorndike (62) who characterized as "erroneous" the "doctrine that some one function is shared by all intellectual traits," and emphasized instead the "singularity and relative independence of every mental process, the thoroughgoing specialization of the mind." This extreme specificity was modified into an implication of group factors when, in 1921, Thorndike (60) stated that "the ability measured by verbal tests is not the same as the ability measured by non-verbal tests. The intelligence test does not give us a measure of some general power which resides in the individual and determines his ability in every variety of intellectual task."

In the same year, Thorndike (63) criticized the available data as being too complicated by large P.E.'s for the intercorrelations to be of service. By means of the inter-columnar correlation (a development of the hierarchy-criterion) Thorndike analyzed the results of 800 soldiers on the 15 sub-tests of Army Alpha and Beta. These results did not substantiate the Two-Factor Theory; neither, however, did they suggest an extreme specificity. Rather is the acceptance of group factors implied in Thorndike's suggestion of a "'number as content' factor and a 'spatial relations as content' factor which act in a fashion midway between nearly complete generality and nearly absolute specificness."

C. M. Davey (10), a pupil of Spearman's, stated as her problem the investigation of the conflicting views of Thorndike. She tested 243 London school-children ranging in age from 8 to 14 years. Verbal tests of same-opposite, synonyms, inference and likeliness; pictorial tests of picture-sequence and enumeration (of items in a picture); and both verbal and pictorial forms of classification, analogies, questions and completion tests were used.[6] Half of the resulting tetrads fell within 1 P.E. of the median, but their average value was .02 and not the zero demanded by a single general factor. Those tetrads exceeding 5 P.E. were then further analyzed and an "intrinsic group factor" asserted as present in four of the verbal tests (same-opposite, synonyms, classification and questions). This "intrinsic group factor" was attributed rather to the content of the tests than to a factor of "verbality," since the other four verbal tests showed no such factor. Such influences as age, sex, race and social status were considered as constituting an "extrinsic group factor" causing the remaining reliable deviations of the tetrad-difference from zero. Davey's general conclusion was that "a verbal mental test measures the same general factor "g" as does a test similar in form but non-verbal in material, *e.g.*, pictures."

---

[6] Spatial tests were rejected for "these showed themselves to depend for their solution on some specific factor besides the general factor with which we are chiefly concerned."

The most elaborate investigation into the presence, location and extent of group factors to date is to be found in the work of T. L. Kelley (33). Kelley tested 140 seventh-grade children, the average age being 12.94 years, upon speed and power forms of a reading test, arithmetic and manipulation of spatial relationships tests and upon tests of memory for verbal, numerical and symbolic materials. With a particular interest in the differentiability of the bonds present among his tests, Kelley devised and utilized a method of analysis which allowed a determination of the strength of these bonds. He concluded that his data showed the presence of a verbal, a number, a memory, a mental manipulation of spatial relations, and a speed factor, as well as a general factor, "alpha." This general factor was attributed to the heterogeneity of the subjects in maturity, sex, and general nurture. Similar results were obtained for a group of 110 third-grade children.

H. E. Stead used in addition to measures of scholastic ability and character, such mental tests as opposites, analogies, synonyms, instructions (directions), completions, absurdities, number series and cipher (decoding of words and sentences in which a, e, i, o, u, h, were represented by punctuation marks) (54). The mental test intercorrelations, reported for 125 boys ranging in age from 12 to 13 years (127 Elementary School and 8 Secondary School pupils), range from .31 ± .037 to .58 ± .057. Applying Spearman's hierarchy criterion to his correlations, Stead concluded that there was clear evidence of a general factor among these mental tests.

Kelley has reported the results of a tetrad-analysis of these data by McCormack (33). He interprets this analysis to indicate an apparent bond between the cipher-number-series tests (both involving a search for missing elements), and weaker bonds between the cipher-absurdities and synonym-antonym tests. The last bond is probably verbal and may involve thinking in terms of antithesis.

Schneck (48) tested 210 New York City College students, all males, of an average age of 235.0 ± 11.3 months with five tests of verbal ability—vocabulary, opposites, analogies, sentence-completion, and disarranged sentences; four tests of number ability—arithmetic reasoning, number series, equation relations and mental multiplication; and four memory tests—picture-picture, word-word, word-nonsense-syllable, geometrical forms-digits. The average intercorrelation of the verbal tests was .5238; the average intercorrelation of the numerical tests .3507; and the correlation between the verbal and numerical tests .1223. When the tetrad-difference criterion was applied to the verbal and number tests separately, evidence was obtained indicating a group factor of *verbality,* V, and a group factor of *numerical ability,* N. These two factors correlated .2625, indicating a high degree of independence. Schneck concluded that V and N do not measure the same ability and that each is distinguishable from retentivity as measured by his memory tests.

Peatman (42) has analyzed the sub-test groupings (reading-comprehension, verbal elements, numerical elements, directions and information) of the Thorndike Entrance Examination from results for 568 Columbia College candidates. A preliminary tetrad-analysis gave presumptive evidence for a group factor common to the Reading Comprehension and Verbal Ability tests. When these were pooled, the resulting tetrad set-up indicated a factor common to all the sub-tests, and factors specific to the four test groups. The contributions of the general and specific factors to the variance of each of the four tests were estimated; and the final conclusion reached that 75% of the total examination was attributable to the common factor, 10% to the specific factors, and 15% to chance errors of measurement.

Stephenson, a pupil of Spearman's, tested 1,037 English schoolgirls ranging in average age from 10 years 2 months to 12 years 5 months with an extensive battery of 8 non-verbal and 8 verbal subtests (55, 56, 57). The non-verbal tests comprised alphabet-construction, code, shape-fitting, picture completion, form analogies, cube-counting, XO series completion and overlapping shapes. The verbal tests included synonyms, sentence completion, classification, interchanged words, opposites, analogies, always has, and directions. The time allotments ranged from 2½ to 5 minutes with an average of 3.99 minutes.

Separate tetrad analyses were made of the verbal, the non-verbal, and the verbal-non-verbal intercorrelations. A comparison was made for each of these categories of the deviation of the distribution of the observed tetrad-differences from the theoretical distribution. This was done on the basis of the divergence of the observed probable error of the mean tetrads of each category of tests from the theoretical probable error. All such divergences were termed "excess errors." The excess error disappeared for the non-verbal tests when the crude scores were rescaled into a "standard" normal distribution. A "specificality" was still noted between the code and overlapping shapes test, however. In the case of the verbal tests the excess error remained even after rescaling.

The following causes of the excess error were considered: a specificality deriving from the similarity of relations involved in several of the tests; idiosyncracies in the construction of the tests; "speed" influences; propinquity effects (habituation, end-spurt); inequalities in schooling; age; calculation errors; etc. None of these was sufficient singly to produce the excess error, though cumulatively all might have been.

The conclusion finally reached was that the verbal sub-tests contained not only the universal general factor *g* found in both verbal and non-verbal tests, but also a smaller general factor V existing throughout verbal abilities.

5. Summary

From this survey of the most important and relevant research available on our problem it is obvious that a simple two-factor pattern cannot adequately account for the correlations which have been obtained. It has been necessary in almost every case to call upon additional group factors of a verbal, numerical or spatial nature to explain the close relationship existing among certain abilities as differentiated from others. These factors have been variously interpreted as caused by the lack of sufficient discreteness of the abilities tested (51, 52), intrinsic group factors of content or extrinsic group factors of age, sex, race and social status (10), and "specificalities" (55, 56, 57) by the proponents of the two-factor theory (*cf.* Tryon, 65, 66). Other workers in this field have been content to locate the factors or to describe them in terms of the test-materials used without attempting to explain them away (4, 33, 37, 48). It is an interesting commentary on the lack of agreement prevailing in this field that the same influences, viz., age, sex, and social status (general nurture) are called upon to explain both the presence of group factors (Davey, 10) and the nature of the general factor (*cf.* "alpha" factor, Kelley, 33).

CHAPTER II

TESTING SET–UP

1. Problem

1. Do the sub-tests within a general intelligence battery measure fundamentally the *same* ability, regardless of the *type* of test material employed?

2. *Within* a given category, *i.e.,* verbal, numerical, spatial, etc., are the test abilities more closely related than they are *between* categories, so that the existence of independent and differentiable mental traits may be demonstrated?

2. Tests

Detailed information concerning the twelve tests used is contained in Table II. The immediate aim in the selection of the battery was the most comprehensive sampling possible of those representative verbal, numerical and spatial abilities commonly measured by intelligence tests. The criteria employed were similar to those formulated by Kelley (33):

1. The variables studied were many and reasonably discrete. Included were three numerical, four verbal and four spatial group tests, and an individual sub-battery of four ''Performance'' tests.

2. All measurements, except perhaps the Goodenough, were ''defined in terms of objective scores on designated tasks.''

3. As far as possible tests of known high reliability were chosen.

Forty minutes were allowed for the language tests (except in the reading test). Very few subjects, although many attempted them, had any of the last ten items correct. A similar situation prevailed for some of the Otis subjects, most of the Beta sub-tests and all of the International, which was the most advanced test included in the battery. In addition to the tests, a questionnaire (Appendix I) was submitted to each child. The younger and duller children and any who needed help in understanding the questionnaire were interviewed individually and those questions which they were unable to answer copied for their parents to complete. The dates of birth were obtained from the school records.

Because of the possibility that the greater amounts of bilingualism in the experience of some of the children (5, 6, 8, 26, 35, 39, 44, 68) might tend to affect unfavorably their scores in tests of verbal-

TABLE II
Description of Tests

Test	*Source*	*Nature of Adaptation*	*Sub-Tests*	*No. of Items*	*Time*	*Graded?*	*Scoring*	*Reliability*	*Given by*
1. Number Series	Miscellaneous	Items from Shakow-Kent, C.U., Terman, Army Alpha, graded according to pretest 50 10-year old S's.	None	40	40 Min.	Yes	No. Right	Split-Halves Correlation + Spearman-Brown Correction.	Teacher
2. Arithmetic Reasoning	I.E.R. CAVD 1925	Levels E–J. Words causing reading difficulty explained to children.	None	60	40 Min.	Yes	No. Right	“ “ “	“
3. Computation	Buswell-John Diag. Chart in Fund. Processes	Ten minute time limits for sub-tests. Regradation for division to correspond with N.Y.C. arithmetic syllabus.	1. Addition	46	10 Min.	Yes	No. Right	“ “ “	“
			2. Subtraction	44	10 Min.				
			3. Multiplication	44	10 Min.				
			4. Division	42	10 Min.				
			Total	176	40 Min.				
4. Vocabulary	I.E.R. CAVD 1925	Levels E–K.	None	70	40 Min.	Yes	No. Right	“ “ “	“
5. Analogies	N.I.T.	60 Items from forms A, B. 10 Items from form C.	None	70	40 Min.	Yes	No. Right	“ “ “	“
6. Sentence Completion	I.E.R. CAVD 1925	Levels E–K. Difficult words read to class. Score-Number Right.	None	70	40 Min.	Yes	No. Right	“ “ “	“
7. Reading	Gates Silent Reading Grades 3–8 Form B, 1926		A. General Significance	24	6 Min.	No	No. Right	Retest 28 boys, 37 girls, in two ‘‘normal’’ and middle classes, *i.e.*, $3B^3$, $4A^3$, Form A; 3 weeks later.*	
			B. Prediction of Outcome	24	8 Min.	No			
			C. Precise Directions	24	8 Min.	No?			
			D. Note Details	54	8 Min.	No			
			Total	126	30 Min.				

TABLE II—(*Continued*)
Description of Tests

Test	*Source*	*Nature of Adaptation*	*Sub-Tests*	*No. of Items*	*Time*	*Graded?*	*Scoring*	*Reliability*	*Given by*
8. Otis	Primary Group 1929	Omission of pictorial completion and language sub-tests.	1. Directions	16	1′40″	Yes	No. Right	Retest, same classes as above, 30 boys, 39 girls, Form B; 3 weeks later.*	2 Examiners
			2. Associations	40	1′35″	No	Approx. 30% of no. right		
			3. Maze	10	2 Min.	Yes	No. Right squares		
			4. Picture Sequence	7	2′10″	Yes	No. Right		
			5. Similarities	8	2′ 5″	Yes	No. Right		
			Total	81	9′30″				
9. Beta	Group Examination 1918	First two items of each sub-test used for demonstration. Number matching omitted.	1. Mazes	8	2 Min.	Yes	No. Right squares	Retest same classes as above, 27 boys, 42 girls, 3 weeks later.*	2 Examiners
			2. Cube Analysis	14	2′30″	Yes	No. Right		
			3. X–O Series	10	1′45″	Yes	No. Right		
			4. Digit Symbol	84	2 Min.	No	⅓ no. right		
			5. Picture Completion	18	3 Min.	Yes	No. Right		
			6. Geometrical Construction	8	2′30″	Yes	No. Right		
			Total	142	13′45″				
10. International	Group Mental 1926	"Regular" time limits. Book #2 given first, then Book #1. Each book divided into two testing periods.	1. Cubes	23	7 Min.	Yes	No. Right	Two methods. r between practise and test-tests and between split-halves (*cf.* 18), Lowest r's used in correcting for attenuation (practise and test r's).	2 Examiners
			2. Associations	48	8 Min.	Yes	No. Right		
			3. Similarities	47	5 Min.	Yes	No. Right		
			4. Faces	23	6 Min.	Yes	No. Right		
			5. Mazes	16	6 Min.	Yes	No. Right		
			6. Rhythms	88	7 Min.	Yes	No. Right		
			7. Analogies	44	5 Min.	Yes	No. Right		
			8. Narratives	48	8 Min.	Yes	No. Right		
			Total	337	52 Min.				

TABLE II—(*Continued*)

DESCRIPTION OF TESTS

Test	*Source*	*Nature of Adaptation*	*Sub-Tests*	*No. of Items*	*Time*	*Graded?*	*Scoring*	*Reliability*	*Given by*
11. Man-Drawing	Goodenough 1926	None	None	Dependent on ability	Approx. 10 Min.		Points for items present. (*cf.* 32)	Retest smaller sample, 17 boys, 22 girls, 11 weeks later.	Teacher
12. Performance	Pintner-Paterson Scale	None except selection of four tests.	1. Five-figure	Dependent on ability	5 Min.		Time & Errors	None obtained	3 Examiners (+2 Group Examiners)
			2. Casuist		5 Min.		Time & Errors		
			3. Knox Cube			Yes	No. of ''lines''		
			4. Healy A		5 Min.		Time & Errors		

---

* The self-r's obtained for the smaller groups were stepped up with Kelley's formula $\frac{\sigma}{\Sigma}=\frac{\sqrt{1-R}}{\sqrt{1-r}}$ for entire group (*cf.* 17, p. 272).

† A single combined score was computed for the three ''time'' scores, the three ''error'' or ''move'' scores and number of ''lines'' as follows: The original seven distributions were converted into distributions with an average of 50 and an S.D. of 14. The seven converted scores were calculated for each subject, added and used as the single Performance score.

ity, the questions on the amount of Yiddish spoken in the home were included. The treatment and results of these data will be reported in Chapter III.

3. Subjects

In deciding upon the age of the subjects for this study, several factors had to be considered. It was necessary to have a group old enough to be able to read and to be well accustomed to paper-and-pencil situations, but young enough to have been minimally affected by selection and differentiation of interests and background.

Five hundred thirty-seven children in the third and fourth grades of Public School 225, in the Brighton Beach section of Brooklyn, New York, were originally given the tests.[7] Because complete test-data were desired for each subject, and because it was considered inadvisable to complicate the situation by introducing so-called ''racial'' differences, only those who had taken all of the tests and were of Jewish parentage were included in the final test-group. Fifty-nine children were excluded because of absence, and 83 because of non-Jewish parentage. One hundred eighty-nine boys and 206 girls remained in the final test group. In the case of International all results are based on 161 boys and 186 girls. This test was given in four separate testing periods because of its length, so that there were more absentees than in other tests.

The final test-group, with the exception of a few who had immigrated in early infancy, were American-born, and had for the most part been born and received their schooling in New York City. The average age was for the boys 9.15 ± .74 years, and for the girls 9.04 ± .71 years.

From Table III it is evident that about an equal number of the parents were born in Russia and in the United States, decidedly fewer in Austria and Poland, and the greater part of the remainder in other European countries. Of the parents not born in the United States, most had been here over 15 years, or long enough in all probability to have had some elementary schooling, or at least to have become literate. Reliable information concerning the education of the parents is unfortunately not available. A disproportionately large percentage, in view of the small percentage of pro-

---

[7] The writer wishes to thank Mr. Solomon Branower, principal of Public School 225, Brooklyn, New York, for permission to conduct the testing. The generous cooperation of the teachers of the third and fourth grades is greatly appreciated.

TABLE III
Description of Subjects

Country	Father's Birthplace Boys	Father's Birthplace Girls	Mother's Birthplace Boys	Mother's Birthplace Girls
U. S. A.	64	61	76	71
Russia	69	72	65	76
Austria	13	32	9	25
Poland	19	14	15	13
Roumania	6	6	2	3
Hungary	2	2	4	
England	6	4	7	5
Germany	4	1	2	
Palestine		1		1
Egypt	2	1	1	1
Turkey	3	2	1	2
Canada			1	
Czecho-Slovakia	1			
Bohemia			1	
France		1	1	
Scotland		1		
Capetown				1
Uncertain		8	4	8
Total	189	206	189	206

Years	Years in U. S. A. Fathers Boys	Years in U. S. A. Fathers Girls	Years in U. S. A. Mothers Boys	Years in U. S. A. Mothers Girls
0– 4	1	4	3	6
5– 9	4	6	4	3
10–14	11	7	11	8
15–19	18	22	20	27
20–24	28	33	23	36
25–29	21	36	29	30
30–34	15	14	15	8
35–39	9	6	1	5
40–44	5	3		
Uncertain	13	14	7	12
Total	125	145	106	135

	Age (*in weeks*) Boys	Age (*in weeks*) Girls
Range	404–573	384–563
Mean	475.6931	470.0243
S.D.	38.3578	36.7239
Median	479.06	469.524
V*	8.0636	7.8132
Reliability of Difference	1.4975	
Chances in 100	85	

Father's Occupation	Boys	Girls
Professional	9	12
Business-Owners	64	76
Business-Managers	11	5
Clerical	7	3
Sales	25	28
Skilled Labor	48	62
Semi- and Unskilled	17	5
Uncertain	8	15
Total	189	206

---

* Coefficient of Variation, $\frac{100\ \text{S.D.}}{\text{Mean}}$

fessionals, reported secondary school and college graduation. Because of the pressure of time and the possibility of arousing antagonism, the question was not further urged and the obtained information not utilized.

From the socio-economic viewpoint, the group was definitely middle-class. Both the professional and the unskilled labor classifications were scantily represented, the greater proportion being either skilled laborers, most often in the clothing industries, or owners of small retail establishments. Some of the children were known by the teachers to have quite well-to-do fathers; about the same number of children reported unemployment.

4. Testing Procedure

A. *Order of Tests*

The tests were given from January 23 to February 5, 1932, from 10:30 A. M. to 3:50 P. M. This was the interval following final examinations for the fall term, and as very little school work remained to be taught, the authorities granted permission for the testing to be performed at that time. The testing schedule is reproduced in Appendix II. Verbal and numerical tests were alternated.

When the thirteen classes involved in the testing reached their rooms at 10:30, they were given the verbal or numerical test for that day. Two trained examiners then proceeded from room to room, giving first the Otis, because it seemed easier than the other non-language tests, then Beta, and finally International. As far as possible, these three tests, as well as the Performance, were given to the youngest and the dullest classes first, so that a minimum amount of information would reach those who had not yet taken a particular test.

The three Performance examiners were each given plainly furnished offices and screens so that privacy and lack of distraction were assured. As classes in the New York City public schools are grouped as "1" bright, "2" normal and "3" dull in academic achievement, the children in each class were rotated among the Performance examiners so that no examiner had children of only one group.

B. *Administration of Tests*

The teachers were trained to administer the language tests as these were not difficult to give once uniform instructions were issued. Moreover, it would have been disturbing to the children to be confronted with too many strangers in the course of the extensive testing program. Such matters as accurate timing, (watches were provided), precautions against cheating, a sufficiency of pencils, refraining from indicating clues, the full explanation of all sample problems, the limitation of the motivation to the writer's formulation, and the fact that the teacher's own work with the children was in no way being tested were emphasized in instructing the teachers. These instructions, under the close supervision of the writer, they followed carefully. The examiners who administered the non-language tests had had adequate training and experience. In short, every effort was exerted to maintain good experimental conditions and uniform testing procedure.

C. *Motivation and Rapport*

The motivation for each of the language tests, which were referred to as word and number games, was formulated by the writer and read by the teachers to their classes. Two points were made clear to the children; they were to do their very best and to do as many items as they could. The teacher's attitude was one of encouragement during the test and approval afterwards. In the case of the non-language tests, the motivation provided in the manuals, was, of course, used.

No incentives were employed. Not only was the change from routine school work, especially after the previous week's finals welcomed and enjoyed by the children, but the novelty and diversity of the test materials served amply to sustain the children's eager interest and effort.

CHAPTER III

PRELIMINARY STATISTICAL RESULTS

1. Scoring

The methods used in computing the test scores are described in Table II. The language tests were scored by an assistant who had had intensive experience in the scoring of intelligence tests. Stencils were prepared for these tests and for the International, Otis and Beta. The International was scored by a graduate student in psychology as part of her work upon a master's thesis (45). Otis and Beta was scored by one of the group examiners, and the other group examiner who had had previous experience in scoring this test scored all the Goodenough tests.

2. Reliability Coefficients

The methods used in obtaining test reliabilities are indicated in Table II, and the coefficients listed in Table IV. These r's compare favorably with the reliabilities reported by the test authors (11, 20, 21, 24, 25, 64, 73). For the language tests the range is from .9189 to .9819. The non-language test reliability coefficients range from .5420 for the Beta to .8737 for the International. Had time permitted the giving of Otis and Beta sooner after the first testing, and the use of larger groups, it is conceivable that the reliabilities might have been higher. Yet it seemed of doubtful value to give Otis and Beta too soon after so long, and, to some extent, so similar a test as the International. On the other hand, in a much longer interval growth changes might have occurred.

The writer has been unable to find either in the manuals or in the literature any data concerning the reliability of the Otis Primary (40) or Army Beta (73); or for most of the separate tests in the Performance Scale (23, 31, 43, 47, 59, 72, 73). It is probable, of course, that by eliminating some of the Otis and Beta sub-tests (as described in Table II) in order to decrease the amount of similar content and to eliminate all verbal and numerical materials, the resulting shortness of these tests tended to decrease the reliability. But had they been given in their entirety the overlapping materials would have made the tests almost identical.

It did not seem profitable to attempt to ascertain the reliability of the Performance series by retest. Several children who were

TABLE IV

PRELIMINARY STATISTICAL ANALYSIS OF TEST RESULTS

Test	*Sex*	*Reliability*	*Range*	*Mean*	*S.D.*	*Median*	Sk/σ_{SK}	*V*	*Reliability of sex differences*	
									$D/S.D._{diff.}$	*Chances in 100*
Number Series	B	.9231	1–37	18.8360	8.3431	19.6800	2.4218	44.2934	1.4029	86
	G	.9412	1–34	17.5631	7.3520	18.5190	1.0671	41.8605		
Arithmetic Reasoning	B	.9274	10–55	40.3862	8.6909	42.8190	3.2488	21.5195	4.1715	99.7
	G	.9308	13–55	35.8058	8.3319	37.7780	.8657	22.6375		
Computation	B	.9819	32–152	101.5767	27.3789	103.7100	1.1791	26.9539	.6911	68
	G	.9783	32–151	102.1699	27.2021	102.5930	.2935	26.6244		
Vocabulary	B	.9444	5–65	39.2646	12.0137	41.0155	2.6557	30.5968	.3082	50
	G	.9365	1–63	38.9078	10.9138	40.8140	2.3842	28.0504		
Analogies	B	.9222	9–63	33.5291	12.8954	31.9740	–3.1263	38.4603	.8061	71
	G	.9189	1–62	32.5000	12.4100	29.8890	–5.5415	38.1846		
Sentence Completion	B	.9615	0–60	34.2857	12.3294	34.9140	.1998	35.9608	.9532	74
	G	.9473	9–61	35.4175	11.2024	35.9380	1.4748	31.6296		
Reading	B	.9404	2–100	44.1376	21.9108	42.3610	–1.8203	49.6420	.3054	58
	G	.9580	7–100	44.7961	20.8186	42.5000	–1.9416	46.4741		
Otis	B	.5743	17–52	38.4921	6.5460	39.1980	.4620	17.0061	3.4058	98.9
	G	.6548	5–50	36.0825	7.5092	37.2450	1.3751	20.8112		
Beta	B	.6511	12.7–60.0	34.1328	8.3766	34.0490	– .3053	24.5412	4.6301	99.8
	G	.5420	8.7–54.0	30.6582	7.6475	30.0100	– .5463	24.9444		
International	B	.8737*	100–295	193.9876	45.5429	193.4380	– .8882	23.4772	2.1041	92
	G	.8139	83–186	183.7312	44.9783	180.0010	– .6931	24.4805		
Goodenough	B	.6836	9–42	20.7037	5.9297	21.0080	– .4127	28.6408	3.0896	98
	G	.8050	10–48	22.6165	6.3578	22.2950	–2.2431	28.1909		
Performance	B		231.3–516.9	347.5058	60.8312	341.5000	–2.0941	17.5051	1.9766	91
	G		234.8–569.3	359.3175	57.7153	355.3340	– .0971	16.0625		

* r between odd and even items of whole test. The correlations between the 8 practise tests and the 8 "test" tests were higher, *i.e.*, .9736 for the boys, .9744 for the girls, after the Brown-Spearman correction.

inadvertently given the series twice showed great gains. Evidently there is a large practice element in these tasks which probably acts to penalize the brighter children who work nearer their final limits at the first presentation of the tasks.

The reliability of the Performance scale and of the Otis Primary and Army Beta should be further investigated. In the case of the Performance the optimum time interval which will yield a reliable self-correlation by the retest method is uncertain—either the interval may be too short so that practice effects enter, or too long so that growth changes occur. If, however, the difficulty of the individual tasks were determined, either on the basis of their decreasing difficulty with age or of the difficulty of all the items for any one age group, and the tasks rearranged in the scale in order of difficulty, the correlation of the odds and evens might then be found to indicate an acceptable reliability.

As it is, the items which were selected for inclusion in our Performance series principally for their discriminatory value for the age group used gave the intercorrelations listed in Appendix III. The highest correlations are those between Time and Moves for any one Performance task. Between the Times for any two tasks, or the Moves for any two tasks, the correlations are intermediate in value, while as between the Time for one task, and the Moves for another, they are still lower. All of the correlations involving the Knox Cube Test are insignificant.

Undoubtedly a study is needed of the reliability of Otis, Beta and Performance before they are used for test-research purposes.

3. Distributions

Frequency distributions for each of the tests have been included in Appendix IV. The skewness of each of these distributions was computed by means of the formula $\text{Sk.} = \text{Mdn.} - \frac{1}{2}\,(\text{P}_{.90} + \text{P}_{.10})$, and standard errors from the formula $\text{S.D.}_{\text{Sk.}} = .5185\,\frac{\text{D}}{\text{N}}$ (12). The reliability of the skewness was determined by $\frac{\text{Sk.}}{\text{S.D.}_{\text{Sk.}}}$. Judged by this criterion, three tests yield distributions which are reliably skewed; *i.e.*, Arithmetic Reasoning is positively skewed for the boys, and Analogies is negatively skewed for both boys and girls. These results may be interpreted as demonstrating that the Arithmetic Reasoning is too easy for the boys, while the Analogies test is too difficult for them and much more so for the girls.

4. Sex Differences

A. *Ability*

The boys have higher averages in the Number Series, Arithmetic Reasoning, Vocabulary, Analogies, Otis, Beta, International, and Performance. The girls score higher in the Sentence Completion, Reading, Computation, and Goodenough. Few of these differences are reliable. The Arithmetic Reasoning and Beta yield the most reliable differences, with the boys superior. The boys are reliably superior also, but to a lesser extent in the Otis. The girls are just reliably superior in the Goodenough. Further than these the differences are negligible.

With due regard for the special characteristics of our subjects and tests, these findings offer an interesting basis for comparison with other results on sex differences. Pintner, in his survey of fourteen extensive studies involving omnibus or general group intelligence tests, concludes that they ''fit roughly into the general picture of the slight superiority of the female in verbal materials whether due to early conditioning or to an inherent difference between the sexes in mental traits.'' (44, Chapter 23.) From his analysis of special test results, Pintner concludes that, ''In number or arithmetical tests, boys are frequently found to be superior, whereas in language tests girls are generally superior. . . .'' (*Ibid.*)

In the light of Pintner's conclusion it may be stated for our data that:

1. There are no reliable sex differences in the ability to handle verbal concepts.

2. The only numerical test in which there is a reliable difference is the Arithmetic Reasoning. Pintner mentions the earlier maturity and ''greater docility of the girls in the educational régime'' in accounting for their superiority in tests affected by educational achievement. While the boys and girls in this study are all quite young, the girls are somewhat younger than the boys. It would be quite difficult to determine whether at this early age, boys in their extra-school activities have had much more occasion to deal with numbers than have girls. As far as that interpretation which involves differing inherent capacities is concerned, still another problem confronts us: How does it come about that of our three numerical tests only the Arithmetic Reasoning yields a reliable sex difference? These tests are as discrete as it is possible

to have them with such young children but nevertheless they intercorrelate .6392 for the boys and .6028 for the girls. Chapter V presents evidence pointing to the existence of a number factor. To what extent sex differences in the degree of this factor are attributable to heredity, and to what degree to environment cannot for the present be determined owing to a complexity in the environmental situation which the ambiguity of the statistical results only serves to emphasize.

3. As far as the non-language tests are concerned, boys are distinctly superior in Beta (99.8 chances in 100 of a true difference) and somewhat less so in Otis (98.9 chances in 100). Nevertheless, in so searching a measure of intelligence as the International, the difference is not reliable. This discrepancy will be further discussed in the correlational analysis of Chapter IV.

4. For the Man-Drawing test, the reliable difference in favor of the girls is in accord with Goodenough's own results for this age group (25, p. 82).

5. Kelley, (33, p. 110) postulates a factor for "the manipulation of spatial relationships of the nature 'boys superior to girls.' " His assumption is not borne out by our data inasmuch as the boys, although superior in the Performance series, (which most nearly resembles his spatial tests) are not reliably so. This lack of a reliable superiority of the boys obtains despite the fact that the rôle of language is of little importance in these form-board tests. Our results are in accord with the principle used by Pintner and Paterson (43) in constructing the scale without separating the sexes; and with the lack of a reliable difference in the Paper Form-Board battery of the Minnesota Mechanical Aptitude Tests (41, p. 275) as between boys and girls in the seventh grade. These same seventh grade girls, however, just reliably excelled the boys in the Minnesota Spatial Relations Test. Divergencies in the particular tests, and in the age ranges of the subjects, should be considered.

B. *Variability*

The range and the coefficient of variation $V = \frac{100 \text{ S.D.}}{\text{Mean}}$ were employed as measures of variability (17). The range of scores is greater for the girls in Vocabulary, Analogies, Otis, Goodenough, and Performance. Most of the differences in range are quite small, however, the largest occurring in the Performance.

When the coefficient of variation was used as the criterion of variability the boys were found to be more variable in the Arithmetic Reasoning, Otis, Beta, and Performance. The differences may again be considered as negligible, however, as the largest is the one for Sentence Completion where V for the boys is 35.9608 and for the girls 31.6296.

These findings agree with the conclusion of Winsor based on a survey of the literature that, ''Girls are as variable as boys whenever a large enough sample of a given age has been tested'' (71). Pintner, however, favors the ''the theory of a slightly greater variability in the male sex'' (44).

5. Differential Bilingualism

According to the nature of the child's answers to the ten questions on the amount of bilingualism in his extra-school experience, his verbal test scores were placed in one of two categories; a ''mostly-English-speaking'' or a ''mostly-Yiddish-speaking'' group. The configurations of answers determining the classification follow:

a. Mostly English-speaking
    Only Yiddish spoken is by *one* of following:
        1. Grandparents to subject
        2. Parents to subject occasionally
        3. Parents with each other most of time
        4. Grandparents speaking both Yiddish and English to subject
        5. Parents speaking Yiddish mostly with each other and occasionally with subject

b. Mostly Yiddish-speaking
    Any other combinations

These groupings are admittedly neither definitively quantitative nor objective. As far as the former is concerned, it is difficult to conceive of an adequate method of determining a numerical score; as regards the former, an attempt was made to objectify the scale by submitting the questions to experts in psychology to be rated for importance. On the basis of these ratings and the writer's own knowledge of the children's home life, the above classifications were determined.

The reliability of the differences in the verbal test averages between the two linguistic groups are reported in Table V.

It is obvious that these Jewish children are not penalized on English verbal tests by being subjected to a greater amount of non-English speech.

TABLE V

THE RELIABILITY OF THE DIFFERENCES BETWEEN THE AVERAGES OF THE BILINGUAL GROUPS ON THE VERBAL TESTS*

	Diff.	$\frac{Diff.}{S.D._{diff.}}$	*Chances in 100*	*Diff.*	$\frac{Diff.}{S.D._{diff.}}$	*Chances in 100*
Vocabulary	1.0300	.5771	73	.9349	.5846	73
Analogies	.1246	.0648	52	1.6900	.9500	83
Sent. Compl.	2.0923	1.1588	87	.9859	.6138	73
Reading	4.6923	1.3965	92	1.2191	.4080	65

* All differences favor the mostly-English speaking group.

These results conflict with those obtained for non-Jewish bilinguals in research reported by Pintner (43, Chapt. 25), Colvin (8) and Klineberg (35) when Welsh, Indian, Italian, etc., were the other languages involved. They are in accord with another study based on a Jewish group (39). Here 149 seventh-grade girls with an average age of 12.88 years and of foreign-born parentage yielded no significant correlations between the proportion of English spoken in the home and scores on such verbal tests as the Otis and the Thorndike Verbal Knowledge. Similar results were obtained with ten-year old girls on the Stanford Binet and Pintner Paterson (26). It may be that the adaptability and assimilability of the Jew as against those of the non-Jew are here in question (*cf.* Pintner, 43, p. 455).

CHAPTER IV

ANALYSIS OF CORRELATIONAL RESULTS[8]

1. Preliminary Treatment of Correlations

The original correlations for boys and girls are reproduced in Table VI. The r's corrected for attenuation are shown in Table VII; and the correlations with the influence of age variability held constant are reproduced in Table VIII.

When corrected for attenuation, the coefficients obtained for those tests of high reliability showed, of course, a negligible amount of change. The correction of the less reliable tests, however, such as Otis, Beta or Goodenough produced an appreciable amount of change. In several tests of low reliability, the correction was so large, *e.g.*, a change of from .5932 to 1.1044 in the Otis-Beta correlation for the girls, as utterly to distort the initial values. In view of the fact that these changes were either inconsiderable or of such magnitude as to produce apparent artifacts, it seemed advisable to work with the actual coefficients obtained. Subsequent statistical analyses, therefore, are based solely upon the raw correlations.[9]

The partialling out of age produced only minor modifications—the highest change for the boys was one of .0637 and for the girls .0723, both occurring in the correlation between Computation and Beta. The correlations of the tests with age are in general the lowest. Had a less discriminating measure been used, such as months, undoubtedly the age correlations would have been still lower, as the scores would have been bunched closer together. It is possible that since the Computation tests an ability in exactly the form in which it is taught in school, and that since the older children who have been in school longer will have had longer drill, the high correlations between age and Computation are to be expected.

---

[8] The means, standard deviations and raw correlations were computed by the Columbia Statistical Bureau. All other computations were made by the writer and checked by her or an assistant.

[9] The use of corrected coefficients in the tetrad-analysis would have resulted in larger tetrad-differences (*cf.* Anastasi, 1, p. 45).

TABLE VI

ORIGINAL CORRELATIONS

Above Diagonal: Boys N = 189
Below Diagonal: Girls N = 206

Variable	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*
1. Age		.1090	.1790	.4247	.1398	.1565	.1609	.1531	.1389	.2728	.2182	.1858	.1407†
2. Number Series	.1412		.6929	.5871	.5285	.6118	.6135	.4421	.4947	.4586	.5829	.3618	.2644
3. Arithmetic Reasoning	.1875	.6329		.6376	.7270	.7216	.7834	.6763	.4578	.3758	.5880	.3129	.2314
4. Computation	.4305	.5558	.6197		.5432	.6040	.6431	.5679	.4436	.4681	.5401	.2143	.2798
5. Vocabulary	.2498	.4881	.6285	.6242		.7036	.7535	.7600	.3953	.2909	.5322	.3129	.1625
6. Analogies	.2441	.5539	.7005	.5620	.6810		.7435	.6809	.4516	.3852	.6154	.2856	.2918
7. Sentence Completion	.3103	.5709	.6979	.6544	.7647	.7007		.7611	.4670	.3812	.5722	.3197	.2644
8. Reading	.2110	.4787	.6062	.5935	.6899	.6670	.7877		.4073	.3167	.5414	.2297	.1301
9. Otis	.2255	.4594	.5266	.3970	.4752	.5772	.5769	.5821		.5932	.6019	.4179	.4038
10. Beta	.3353	.5184	.5081	.5532	.3915	.5014	.4712	.4905	.6579		.6106	.4322	.4463
11. International*	.2026	.5801	.6347	.5134	.5522	.6688	.6245	.4701	.6798	.5827		.3473	.3428
12. Goodenough	.1491	.3500	.3749	.2777	.3601	.3744	.4417	.3058	.3795	.3004	.3810		.3255
13. Performance	.0622	.3953	.4731	.2654	.3113	.3295	.3808	.2843	.4195	.3267	.2991	.2719	

* N, Boys = 161, Girls = 186.

† The positive correlations with Performance indicate that the lower scores in this test which represent the greater degrees of ability in it (since it was scored for time and errors) tended to be positively associated with high scores in the other variables.

TABLE VII

Intercorrelations Corrected for Attenuation

Above Diagonal: Boys N = 189
Below Diagonal: Girls N = 206

Variable	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
1. Age		.1134	.1859	.4296	.1439	.1630	.1641	.1579	.1833	.3381	.2334	.2247
2. Number Series	.1455		.7488	.6166	.5660	.6631	.6511	.4745	.6794	.5915	.6490	.4511
3. Arithmetic Reasoning	.1943	.6762		.6682	.7769	.7803	.8296	.7242	.6273	.4837	.6533	.3923
4. Computation	.4352	.5792	.6494		.5641	.6347	.6618	.5910	.5908	.5854	.5831	.2620
5. Vocabulary	.2581	.5199	.6732	.6521		.7540	.7907	.8065	.5368	.3710	.5859	.3911
6. Analogies	.2546	.5956	.7574	.5927	.7342		.7895	.7312	.6206	.4971	.6856	.3603
7. Sentence Completion	.3188	.6046	.7432	.6798	.8119	.7510		.8004	.6284	.4818	.6243	.3951
8. Reading	.2156	.5041	.6420	.6131	.7284	.7109	.8268		.5543	.4047	.5973	.2838
9. Otis	.2787	.5851	.6745	.4960	.6068	.7441	.7325	.7350		.9701	.8498	.6247
10. Beta	.4554	.7257	.7153	.7597	.5496	.7105	.6576	.6807	1.1044		.8096	.7100
11. International*	.2246	.6627	.7292	.5753	.6325	.7734	.7112	.5323	.9311	.8773		.4656
12. Goodenough	.1662	.4060	.4339	.3124	.4130	.4345	.5020	.3515	.5582	.4150	.4543	

* N, Boys = 161, Girls = 186.

TABLE VIII

INTERCORRELATIONS WITH AGE VARIABILITY HELD CONSTANT

Above Diagonal: Boys N = 189
Below Diagonal: Girls N = 206

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1. Number Series		.6886	.6010	.5215	.6057	.6075	.4331	.4872	.4485	.5764	.3496	.2531
2. Arithmetic Reasoning	.6235		.6306	.7206	.7138	.7771	.6675	.4443	.3455	.5717	.2892	.2117
3. Computation	.5539	.6079		.5397	.6010	.6433	.5622	.4290	.4044	.5064	.1522	.2455
4. Vocabulary	.4724	.6115	.5912		.6970	.7480	.7548	.3833	.2654	.5192	.2949	.1457
5. Analogies	.5410	.6873	.5220	.6602		.7368	.6731	.4395	.3604	.6031	.2643	.2759
6. Sentence Completion	.5601	.6851	.6070	.7466	.6779		.7551	.4541	.3552	.5576	.2988	.2474
7. Reading	.4639	.5901	.5698	.6732	.6493	.7772		.3944	.2892	.5268	.2073	.1110
8. Otis	.4433	.5061	.3411	.4441	.5527	.5473	.5613		.5828	.5915	.4029	.3919
9. Beta	.5051	.4811	.4809	.3373	.4593	.4100	.4559	.6345		.5870	.4035	.4282
10. International*	.5688	.6203	.4822	.5289	.6521	.6033	.4465	.6647	.5580		.3199	.3230
11. Goodenough	.3360	.3572	.2392	.3445	.3525	.4206	.2838	.3591	.2688	.3623		.3077
12. Performance	.3912	.4706	.2648	.3061	.3247	.3810	.2780	.4171	.3252	.2931	.2661	

* N, Boys = 161, Girls = 186.

2. PROBABLE ERRORS

In the table which follows are indicated those raw r's for the two sets of N's which are most nearly four times their P.E.'s for the sexes, and therefore just reliable. Thereafter the P.E. decreases as the r increases and greater reliability obtains.[10]

	Boys		*Girls*	
N	*189*	*161*	*206*	*186*
Variables	*Computation–Goodenough*	*Age–International*	*Age–Arith. Reasoning*	*Age–International*
r	.2143	.2182	.1875	.2026
P.E.$_{\cdot r}$	± .0468	± .0560	± .0453	± .0476
$\frac{r}{P.E._{\cdot r}}$	4.5791	4.3123	4.1391	4.2743

For the boys, all correlations between any two tests are reliable except the r's between the Vocabulary-Performance and Reading-Performance tests, while for the girls all test correlations are significant. All correlations with age are unreliable except, for the boys, Computation and Beta; while for the girls the only unreliable correlations with age are those of Number Series, Goodenough and Performance.

3. AVERAGE INTERCORRELATIONS

The correlational averages which are reported in Table IX were obtained by Fisher's method of transforming r's into their corresponding z-values.[11] Age correlates least with the other variables and Sentence Completion the most. Of the test variables those which correlate least with the others are Goodenough and Performance. The average correlation of each test with all the others ranges from .2913 for Performance to .5980 for Sentence Comple-

---

[10] To facilitate the evaluation of the r's appearing in other tables, the minimum coefficients which would be just 4 P.E. for the four populations have been computed. They are for the boys, N = 189, r = .1892, N = 161, r = .2038; for the girls, N = 206, r = .1818, N = 186, r = .1906. These r's were obtained by substituting in the formulae: (1) $\frac{r}{P.E._{\cdot r}} = 4$. (2) $\frac{r}{.6745(1-r^2)/\sqrt{N}}$ and solving the quadratic equations which resulted.

[11] *Cf.* Fisher (14), p. 164. S.D.$_{\cdot r}$ depends on the true value of the correlation ρ, which is unknown. S.D.$_{\cdot z}$ is, however, practically independent of ρ. Moreover, z-values are more nearly constant and more normally distributed than the observed r's. Consequently, "The values of z give a truer picture of the relative importance of correlations of different sizes than do the values of r."

TABLE IX

AVERAGE CORRELATIONS—(Z-VALUES)

Variables	*Boys*	*Girls*
Category Correlations		
4 Verbal Tests	.7352	.7211
3 Numerical Tests	.6411	.6044
5 Non-Language Tests	.4621	.4462
Verbal-Numerical	.6291	.6044
Verbal-Non-Language	.3800	.4699
Numerical-Non-Language	.4136	.4621
Average for all Tests	.5580	.5580
Each Variable With All Others		
Age	.1877	.2355
Number Series	.5227	.5154
Arithmetic Reasoning	.5850	.5915
Computation	.5154	.5370
Vocabulary	.5441	.5580
Analogies	.5717	.5850
Sentence Completion	.5980	.6231
Reading	.5299	.5580
Otis	.4699	.5299
Beta	.4382	.4854
Goodenough	.3275	.3540
International	.5370	.5511
Performance	.2913	.3452
Each Test With Categories		
Vocab. & Other Verbal	.7398	.7163
Anal. " " "	.7114	.6858
Sent. " " "	.7531	.7531
Read. " " "	.7352	.7211
No. Series & other Numerical	.6411	.5980
Each Test With Categories (Cont'd.)		
Arith. Reas. & other Numerical	.6640	.6291
Comput. " " "	.6107	.5915
Otis & other Non-Language	.5080	.5441
Beta " " " "	.5227	.4777
Internat. & other Non-Language	.4854	.5005
Goodenough & " " "	.3800	.3364
Performance & " " "	.3800	.3275
Vocab. & Numerical	.6107	.5857
Anal. " "	.6469	.6107
Sent. " "	.6858	.6411
Read. " "	.5717	.5649
Vocab. & Non-Language	.3452	.4219
Anal. " " "	.4136	.5005
Sent. " " "	.4053	.5005
Read. " " "	.3364	.4382
Otis & Verbal	.4301	.5580
Beta " "	.3452	.4621
Internat. & Verbal	.5649	.5850
Goodenough & "	.2821	.3714
Performance & "	.2165	.3275
No. Series & Non-Language	.4382	.4621
Arith. Reas. & Non-Language	.3969	.5080
Computation " " "	.3969	.4053
Otis & Numerical	.4621	.4621
Beta " "	.4301	.5227
Internat. & Numerical	.5717	.5784
Goodenough & "	.3004	.3364
Performance & "	.2636	.3800
No. Series & Verbal	.5511	.5227
Arith. Reas. & "	.7306	.6584
Computation & "	.5915	.6107

tion for the boys, and from .3452 to .6231 for the same tests for the girls.

Among the tests of each category, the four verbal tests correlate highest among themselves—boys, .7338; girls, .7152; the numerical tests somewhat less—boys, .6392, girls, .6028, and the non-language tests least—boys, .4522, girls, .4299.[12] As concerns the average correlation between any two categories, the verbal-numerical correlation is highest—boys, .6219, girls, .5966. For the boys, the numerical-spatial is lower, .4049 and the verbal-spatial lowest, .3766.[13] For the girls, the average verbal-spatial and numerical-spatial are about the same, .4585 and .4552. The average correlation of all tests is .5580 for both boys and girls. Apparently, such minor discrepancies as are observable between the sexes in Table IX, and which are in most cases very small, reaching in only three instances a magnitude of .12, (Otis-Verbal, Beta-Verbal, Number-Performance) balance each other nicely when the average is taken for the battery as a whole. Invariably the correlations are higher for the girls when the discrepancy is .06 or over.

The tests of each category which correlate highest with all the other tests are the Sentence Completion, *viz.*, .5980 and .6231, Arithmetic Reasoning, *viz.*, .5850 and .5915, and International, *viz.*, .5370 and .5511, for boys and girls, respectively. As between one test and the others in its category, Sentence Completion correlates highest with the other verbal, .7531 and .7531, Arithmetic Reasoning with the other numerical, .6640 and .6291 for boys and girls, respectively. Beta correlates highest with the other spatial tests for the boys, namely, .5227 and Otis for the girls, .5441.

Turning to the correlation between any test in one category with all the tests in any other, we find that the following results hold for both boys and girls: Sentence Completion correlates highest with the numerical tests, .6858 and .6411, Analogies and Sentence Completion about equally with the non-language category, .4136 and .4053, respectively, for the boys, while .5005 is the average for both boys and girls. International is the spatial test which correlates highest with both the verbal and numerical sub-batteries, .5649 and .5850 for the verbal, boys and girls, respectively, and .5717 and .5784 for the numerical. Between any one numerical test and the spatial category, the correlations between the latter and Arithmetic Reasoning and Computation are equal for the boys, .3969, while the Arithmetic Reasoning correlates highest with the spatial tests for the girls, .5080. The number test which correlates highest with the verbal tests is the Arithmetic Reasoning for both sexes; .7306 and .6584 for the boys and girls, respectively.

---

[12] *Cf.* 10, 19, 48, 55, 56, who also report higher correlations among verbal tests than among tests of other categories.

[13] The low average intercorrelations among the spatial tests are probably to some extent attributable to their low reliability. The z-value averages for the corrected coefficients are as follows—Boys: Spatial r = .8076, Numerical-Spatial r = .5511; Girls: Spatial r = .8511, Numerical-Spatial r = .6044.

The relatively high correlations between the verbal and number tests can probably be explained in terms of the aims and practises of elementary school pedagogy. The teaching of arithmetic in the elementary schools emphasizes the speediest possible transformation of numerical relationships into verbal concepts. For instance, the word six and the symbol six are demonstrated concretely as applying to situations in which 3 and 3, 2 and 4, or 1 and 5 are contained. Once the objective relationship is grasped, in all subsequent drills for either accuracy or speed, the child is actually conditioned to substitute a verbal response for the intermediate numerical operations. It is therefore to this verbalization of arithmetical processes, that some part at least of the verbal-numerical correlation is attributable.

In connection with the discussion in Chapter III on the reliability of the sex differences in the spatial tests, we find that the International, which does not given a reliable sex difference, correlates much higher with the verbal and numerical tests for the boys, and somewhat higher for the girls than do the other spatial tests. Evidently, especially for the boys, the abilities involved in solving the tasks of International are much more closely allied to the abilities measured by number and verbal tests than those needed for Otis and Beta. From these results a linguistic inferiority might be argued for the boys.

In any event, the relatively high correlations obtained between International and tests which present actual verbal and numerical concepts, should here be remarked. These latter tests are the more susceptible to the influence of previous training. At this point in anticipation of and dependence upon subsequent statistical analysis, the problem of the interpretation of these results arises. Should we be concerned only with the deduction of the psychological factor organization of these abilities in young children as based on statistical techniques? Or do the results rather suggest that the more fundamental problem is that of the ultimate nature of the reasoning process, its resolution into symbolization, imagery and whatnot, and its identification with *g*, if *g* exists?

It may be contended that the similarity of test materials would provide a simpler basis for the explanation of the above correlations. International contains analogies, narrative sequence (*cf.* Prediction of Outcome sub-test in the Gates Reading Test, or even Sentence Completion), cube-counting, etc. In addition, such factors as similarity of test-administration, of timing, of general paper-and-pencil situation and group-test conditions might also be offered as a simpler explanation of the relatively high correlations.

CHAPTER V

FACTOR ANALYSIS OF CORRELATIONAL RESULTS

1. Tetrads within Each Test-Category

The technique of tetrad analysis was applied to the data in order to determine the existence and nature of common and specific factors or functions throughout the tests. The results of the separate analyses of the verbal, the numerical, and the spatial tests, made in order to determine whether they would yield evidence of a function common to each of the categories, will first be presented. The methods and results of the computation of tetrads across these classifications, for the determination of the differentiability of these functions, will then be discussed.

It may not be amiss here to set down an explanation of nomenclature:

V = a factor of Verbal Ability common to the group of four word-tests.

N = a factor of Numerical Ability common to the group of three numerical tests.

S = a factor comprising the abilities to handle the Spatial materials of the five non-language tests.

C = for any one category, the factor V, C, or S *common* or *central* to it.

A. *The Verbal Tests*

The results for the verbal tests follow:

TABLE X

Verbal-Test Tetrads

1. Vocabulary
2. Analogies
3. Sentence Completion
4. Reading

	Boys			*Girls*		
	t_{1234}	t_{1243}	t_{1342}	t_{1234}	t_{1243}	t_{1342}
td.	.0224	−.0296	−.0520	.0263	.0530	.0267
$P.E._t$*	± .0105	± .0248	± .0199	± .0521	± .0190	± .0144
$\frac{t}{P.E._t}$	2.1333	1.1938	2.6131	1.0478	2.7895	2.3421

* Spearman's formula (51, App. p. x, 39).

The tetrad-difference is in each case well within the 4 P.E. criterion which will be used in this study. A central and specific verbal factors are therefore indicated.

The reliability of these tetrad-differences was further examined by means of Holzinger's formula for the determination of boundary conditions for four variables in which $r_{12}r_{34} = r_{13}r_{24} = r_{13}r_{24} = l$ and $4(r^2_{12} + r^2_{13} + r^2_{14} + r^2_{23} + r^2_{24} + r^2_{34} + r_{12}r_{13}r_{23} + r_{12}r_{14}r_{24} + r_{13}r_{14}r_{34} + r_{23}r_{24}r_{34}) + 3l^2 > 16$ (30, p. 7) for a common factor to be present. For the boys, the left-hand side of the equation equals $20.1303 > 16$, and for the girls, $18.9527 > 16$. The existence of V may therefore be considered as established. Its differentiability will be examined later.

B. *The Number Tests*

Because it had been possible to administer only three tests of a numerical nature, tetrads could not be computed for these tests. Methods of triad analysis and boundary conditions were employed instead.

Kelley's Propositions 4 and 5 were first applied in triad analysis (33). In the former, one common and two specific factors are present "only in case the product of two of the three inter-correlations is equal to the third."[14] As tabulated below, the results do not satisfy this condition:

Boys	*Girls*
$.6929 \times .5871 \neq .6376$	$.6329 \times .5558 \neq .6197$
$.5871 \times .6376 \neq .6929$	$.5558 \times .6197 \neq .6329$
$.6929 \times .6376 \neq .5871$	$.6329 \times .6197 \neq .5558$

Proposition 5 was then tried and checked. First, in order to establish the existence of one common factor and three specifics, "the absolute value of the product of every two of the three inter-correlations" must be less than the third. With the check, the ratio of the product of two of the r's to the third must be less than unity. As indicated below, these conditions were satisfied. The existence of N may thus be considered as established.

	Boys	*Girls*
$r_{xy}r_{yz} < r_{xz} =$	$.6929 \times .6376 = .4418 < .5871$	$.6329 \times .6197 = .3922 < .5558$
$r_{xy}r_{xz} < r_{yz} =$	$.6929 \times .5871 = .4068 < .6376$	$.6329 \times .5558 = .3518 < .6197$
$r_{xz}r_{yz} < r_{xy} =$	$.5871 \times .6376 = .3743 < .6929$	$.5558 \times .6197 = .3444 < .6329$

---

[14] Cureton and Dunlap propose a similar technique for corrected r's (9).

All ratios are less than one.

The boundary condition technique applied was Thompson's. In this method $r^2_{12} + r^2_{23} + r^2_{13} + 2r_{12}r_{23}r_{13} > 1$ for C to be present (61).[15] With x as Number Series, y as Arithmetic Reasoning, and z as Computation; $r_{xy} = .6929$, $r_{xz} = .5871$, and $r_{yz} = .6376$, the left side of the equation is found to be 1.7501 for the boys. This value indicates the existence of a common numerical bond, N, among the three number tests. For the girls, $r_{xy} = .6329$, $r_{xz} = .5558$, $r_{yz} = .6197$; the left-hand side of the equation equals 1.5294 which is again sufficiently greater than 1 to indicate a numerical bond. The existence of N previously established is substantiated by this analysis.

C. *The Spatial Tests*

The tetrad analyses of the non-language tests was complicated by the fact that the N's for the International differed from the N's for the other tests. Only 161 boys and 186 girls took the International. Two additional sets of P.E.'s were therefore calculated for the tetrads in which International was included with the following values for N: for the boys, $\frac{3 \times 189 + 161}{4} = 182$,—this is the average N for the four tests, the N of 189 being given triple weight since it occurs three times; and a P.E. for the minimum value of N = 161, For the girls $\frac{3 \times 206 + 186}{4} = 201$ and the minimum value of N, = 186 were used as N's. Since the larger the N, the smaller the P.E. and the more difficult the satisfaction of the tetrad-difference criterion, the evaluation of the tetrad-differences in terms of the P.E.'s derived from the various N's was essential.

For five variables, fifteen tetrad-differences may be calculated. The range of P.E.'s seems to be approximately from .02 to .03. The results of the spatial-test tetrad-analysis appear in Table XI.

The tetrad-differences arising from the spatial tests do not in any case approximate a magnitude reliably greater than zero. We may therefore assume a central factor, S, common to these tests. As hitherto, the method of boundary conditions has been applied to substantiate the existence of this C. As no formula is at present available for a five-variable problem, those five tetrads were chosen for analysis which were composed of such combinations of four variables as do not occur in any other tetrads. The results of this analysis are listed in Table XII.

---

[15] Holzinger has an extension of this method for negative r's (30).

TABLE XI

SPATIAL TEST TETRADS

1. Otis 3. International
2. Beta 4. Goodenough
5. Performance

		Boys			*Girls*	
Variables	*td*	$P.E._t$	$\frac{t}{P.E._t}$	*td*	$P.E._t$	$\frac{t}{P.E._t}$
1234	−.0541	± .0234*	2.3120	.0465	± .0272†	1.7096
		± .0239	2.2636		± .0275	1.6909
		± .0254	2.1299		± .0286	1.6259
1243	−.0492	± .0231	2.1299	.0296	± .0250	1.1840
		± .0236	2.0847		± .0253	1.1700
		± .0251	1.9602		± .0263	1.1255
1342	.0049	± .0292	.1678	−.0169	± .0218	.7752
		± .0297	.1650		± .0221	.7647
		± .0316	.1551		± .0230	.7348
1245	.0066	± .0274‡	.2409	.0549	± .0288‡	1.9063
1254	.0186	± .0259	.7181	.0529	± .0295	1.7932
1452	.0120	± .0214	.5607	−.0020	± .0155	.1290
1345	.0526	± .0288	1.8264	.0713	± .0273	2.6117
		± .0293	1.7952		± .0277	2.5740
		± .0312	1.6859		± .0288	2.4757
1354	.0557	± .0283	1.9682	.0250	± .0319	.7837
		± .0289	1.9273		± .0323	.7740
		± .0307	1.8143		± .0336	.7440
1453	.0031	± .0175	.1771	−.0463	± .0157	2.9490
		± .0179	.1732		± .0159	2.9119
		± .0190	.1632		± .0165	2.8061
2345	.0506	± .0281	1.8007	.0686	± .0252	2.7222
		± .0286	1.7692		± .0255	2.6902
		± .0304	1.6645		± .0265	2.5887
2354	.0438	± .0290	1.5103	.0339	± .0290	1.1690
		± .0295	1.4847		± .0294	1.1531
		± .0314	1.3949		± .0396	.8561
2453	−.0068	± .0185	.3676	−.0347	± .0158	2.1962
		± .0189	.3598		± .0161	2.1553
		± .0201	.3383		± .0167	2.0778
1235	−.0653	± .0243	2.6872	−.0253	± .0249	1.0161
		± .0248	2.6331		± .0251	1.0080
		± .0264	2.4735		± .0262	.9656
1253	−.0433	± .0221	2.0045	−.0476	± .0232	2.0517
		± .0225	1.9869		± .0235	2.0255
		± .0240	1.8458		± .0245	1.9429
1352	.0220	± .0294	.7483	−.0223	± .0264	.8447
		± .0300	.7333		± .0267	.8352
		± .0319	.6897		± .0278	.8022

* N's = 189, 182, 161 in order.

† N's = 206, 201, 186 in order.

‡ N's = 189, 206 respectively for all 4 tests.

TABLE XII*

VALUES FOR HOLZINGER'S EQUATION 13 FOR 4-VARIABLE BOUNDARY CONDITIONS WITH THE SPATIAL TESTS

Tetrad	*Boys*	*Girls*
t_{1234}	8.4658	8.6100
t_{1235}	8.4458	8.5451
t_{1245}	6.1555	5.0964
t_{1345}	5.3101	5.4643
t_{2345}	5.6424	4.1514

* Subscripts indicate same variables as in previous table.

The tetrads vanish, but the correlations do not satisfy the boundary condition, so that a common factor may or may not be present (30). The tetrad-differences in the preceding table may tentatively be re-examined for an explanation of the doubt which the boundary-condition technique casts upon the existence of S. If we analyze the components of all td.'s of the magnitude .0465 or roughly, .05 or over, we find that in those tetrads, for the boys the r_{13} occurs four times and r_{45} three times while for the girls r_{12} occurs three times and r_{45} four times. Possibly the overlapping indicated as existing between these pairs of variables points to subsidiary group factors just small enough to interfere with the isolation of S.

2. CENTRAL FACTOR CORRELATIONS

Given a tetrad-difference not reliably greater than zero, it is possible to determine the correlation of a test with the factor common to it and the other tests composing the tetrad, by means of the formula $r_{xC} = \sqrt{\frac{r_{xy}r_{xz}}{r_{yz}}}$ (51, App.). When four variables are involved, three determinations of r_{xC} are possible; these were transmuted into Fisher's z-scores and averaged. With r_{xC} known, the residual correlation between the specific factors of any two tests was found by holding constant their correlation with the central factor. The values resulting for the verbal tests follow:

TABLE XIII

CORRELATIONS OF VERBAL TESTS WITH CENTRAL AND SPECIFIC FACTORS

1. Vocabulary	3. Sentence Completion
2. Analogies	4. Reading

Boys			*Girls*		
Original r's	r_{xV}	$r_{xy.V}$	*Original r's*	r_{xV}	$r_{xy.V}$
$r_{12} = .7036$	$r_{1V} = .8668$	$-.0079$	$r_{12} = .6810$	$r_{1v} = .8397$	$.0543$
$r_{13} = .7535$	$r_{2V} = .8144$	$-.0860$	$r_{13} = .7647$	$r_{2v} = .7895$	$-.0054$
$r_{14} = .7600$	$r_{3V} = .8917$	$.0684$	$r_{14} = .6899$	$r_{3V} = .9121$	$-.0693$
$r_{23} = .7435$	$r_{4V} = .8565$	$.0659$	$r_{23} = .7007$	$r_{4V} = .8455$	$-.0771$
$r_{24} = .6809$		$-.0554$	$r_{24} = .6670$		$-.0015$
$r_{34} = .7611$		$.0111$	$r_{34} = .7877$		$.0754$

The correlations of the verbal tests with V are quite high, that of Sentence Completion being the greatest and Analogies the least for both sexes. The correlations between the specifics are entirely insignificant.

The results for the numerical tests are analogous:

TABLE XIV

CORRELATIONS OF NUMERICAL TESTS WITH CENTRAL AND SPECIFIC FACTORS

1. Number Series 2. Arithmetic Reasoning 3. Computation

Boys			*Girls*		
Original r's	r_{xN}	$r_{xy.N}$	*Original r's*	r_{xN}	$r_{xy.N}$
$r_{12} = .6929$	$r_{1N} = .7987$	.0000	$r_{12} = .6329$	$r_{1N} = .7533$	.0003
$r_{13} = .5871$	$r_{2N} = .8675$	.0002	$r_{13} = .5558$	$r_{2N} = .8400$	.0002
$r_{23} = .6376$	$r_{3N} = .7350$	.0000	$r_{23} = .6197$	$r_{3N} = .7377$	.0000

Arithmetic Reasoning shows the highest correlation with N and Computation the least for both boys and girls.

The following correlations of the non-language tests with their central factor represent the average of six determinations.

TABLE XV

CORRELATIONS OF SPATIAL TESTS WITH CENTRAL AND SPECIFIC FACTORS

1. Otis 3. International
2. Beta 4. Goodenough
5. Performance

Boys			*Girls*		
Original r's	r_{xS}	$r_{xy.S}$	*Original r's*	r_{xS}	$r_{xy.S}$
$r_{12} = .5932$	$r_{1S} = .7779$	−.1105	$r_{12} = .6579$	$r_{1S} = .9015$	.0404
$r_{13} = .6019$	$r_{2S} = .8144$	.1176	$r_{13} = .6798$	$r_{2S} = .7163$	−.0348
$r_{14} = .4179$	$r_{3S} = .7064$	−.0102	$r_{14} = .3795$	$r_{3S} = .7648$	−.1535
$r_{15} = .4038$	$r_{4S} = .5441$	−.0262	$r_{15} = .4195$	$r_{4S} = .4854$	−.0292
$r_{23} = .6106$	$r_{5S} = .5370$	.0858	$r_{23} = .5827$	$r_{5S} = .4777$	.0776
$r_{24} = .4322$		−.0224	$r_{24} = .3004$		−.0775
$r_{25} = .4463$		.0184	$r_{25} = .3267$		−.0253
$r_{34} = .3473$		−.0624	$r_{34} = .3810$		.0174
$r_{35} = .3428$		−.0610	$r_{35} = .2991$		−.1170
$r_{45} = .3255$		.0470	$r_{45} = .2719$		−.0521

The residual correlation between certain of the specifics is higher for the non-language tests than the others, and the correlation with the central factor, in general, much lower. For the boys, Beta correlates highest, Otis and International almost equally and less, Performance and Goodenough are quite low. For the girls, Otis correlates highest with the central factor, Goodenough and Performance lowest, while Beta and International are intermediate.

To compare these results with the correlational analysis of the tetrads arising from the spatial tests, we find that for the boys, the largest residual correlation .1176 occurs between Otis and International—the correlation which had given the greatest number of larger tetrad-differences. The other correlations between specifics are smaller or negative and do not jibe with the inter-test bonds estimated on the basis of spatial-test tetrad-differences which were in no case, it should be remembered, reliable.

3. Cross-Category Tetrads

With the establishment of the existence of the central factors V, N, and probably, S, each common to the respective test-categories, two questions remained to be answered: Were these C's similar and overlapping in nature so that what was common to the tests in one category was not differentiable from what was common to the tests in another category? Or were the configurations of abilities represented by the various C's so differentiated that they reliably lacked any community of function as measured by the significance of the tetrad-difference when tetrads composed of a pair of tests of each category were analyzed?

Tetrads of this description were set up from correlations selected in two ways. First, those two tests, from two different categories, which correlated highest with their C's were chosen and paired in tetrads. Then on the assumption that if the C factors were found to be differentiable for the two "best" tests of each category, this should be true of the two "poorest" tests of each category, these tests were also paired in tetrads. It seemed desirable to determine whether in tetrad relationships consisting of tests which correlated least with their respective C's the differences would also be reliably greater than zero and again indicate that the C's had no community of function.

The criterion used in evaluating the tetrads, besides their probable error, was Kelley's 16th Proposition. According to this statement if two tetrads are equal to each other and the third equal to zero, the tetrads "may conceivably have arisen from four variables, x_1, x_2, x_3, x_4 through which was a general factor plus, in addition thereto, a second factor common to x_1 and x_2 and a second factor common to x_3 and x_4." (33, p. 69.)

A. *Verbal-Numerical*

The two "best" tests of V as measured by r_{xV} were for the boys Sentence Completion and Vocabulary, and for the girls Sentence

Completion and Reading. For both sexes, the two best tests of N were Arithmetic Reasoning and Number Series.

TABLE XVI

1. Vocabulary
2. Analogies
3. Sentence Completion
4. Reading
5. Number Series
6. Arithmetic Reasoning
7. Computation

a. Tetrads Arising from Two "Best" Tests of V and N

	Boys			*Girls*		
	t_{1356}	t_{1365}	t_{1563}	t_{3456}	t_{3465}	t_{3564}
td.	.1081	.0761	–.0320	.1524	.1644	.0120
$P.E._t$	±.0251	±.0286	±.0046	±.0310	±.0310	±.0146
$\frac{t}{P.E._t}$	4.3068	2.6608	6.9565	4.9161	5.3032	.8219

b. Tetrads Arising from Two "Poorest" Tests of V and N

	Boys			*Girls*		
	t_{2457}	t_{2475}	t_{2574}	t_{1257}	t_{1275}	t_{1572}
td.	.0524	.1328	.0804	.1042	.0328	–.0714
$P.E._t$	±.0346	±.0236	±.0124	±.0216	±.0319	±.0148
$\frac{t}{P.E._t}$	1.5145	5.6271	6.4839	4.8241	1.0282	4.8243

Two tetrad differences equal to each other, and the third equal to zero are found in only one set-up, the girls' "bests." Here, therefore, the presence of V or N is clearly indicated. In each of the other set-ups one reliably significant difference obtains, but no other two equal each other. A significant bond exists in the boys' "poorests," t_{2574}, either between Analogies and Number Series or between Reading and Computation. As between these pairs, the bond may well be present in the latter, since they both test abilities which are part of the school curriculum. The relations between the other single correlations of a verbal and a number test are negligible or reliably negative—*cf.* girls t_{1572} in which an opposition is indicated between Vocabulary and Number Series or between Analogies and Computation.

B. *Verbal-Spatial*

The tetrads derived from the verbal and spatial tests in Table XVII indicate in three out of four cases the existence of a differentiable central factor, V or S. The exception occurs for the girls' ''poorests'' and is probably attributable to the low correlation between Goodenough and Performance, .2719, rather than the lack of a bond between the two poorest verbal tests as these offer evidence of a bond when combined in tetrads with the number tests.

TABLE XVII

1. Vocabulary
2. Analogies
3. Sentence Completion
4. Reading
5. Otis
6. Beta
7. International*
8. Goodenough
9. Performance

a. Tetrads Arising from Two ''Best'' Tests of V and S

	Boys			*Girls*		
	t_{1356}	t_{1365}	t_{1563}	t_{3457}	t_{3475}	t_{3574}
td.	.2963	.3111	.0148	.2643	.1720	−.0923
$P.E._t$	± .0410	± .0403	± .0160	± .0284 ± .0287 ± .0299	± .0403 ± .0409 ± .0425	± .0198 ± .0200 ± .0208
$\frac{t}{P.E._t}$	7.2268	7.7196	.9250	9.3063 9.2091 8.8395	4.2680 4.2054 4.0471	4.6616 4.6150 4.4375

b. Tetrads Arising from Two ''Poorest'' Tests of V and S

	Boys			*Girls*		
	t_{2489}	t_{2498}	t_{2894}	t_{1289}	t_{1298}	t_{1892}
td.	.1844	.1546	−.0298	.0665	.0686	.0021
$P.E._t$	± .0342	± .0365	± .0087	± .0302	± .0301	± .0132
$\frac{t}{P.E._t}$	5.3918	4.2356	3.4252	2.2020	2.2791	.1591

* Three P.E.'s calculated for International as previously explained.

C. *Numerical-Spatial*

The tetrads from the two ''best'' tests of N and S as indicated in Table XVIII offer evidence for the existence of either or both of these factors according to Kelley's criterion. For the boys, the evidence is less conclusive for the poorest tests, only one of the tetrad-differences—t_{1387}—being reliable while for the girls' ''poorests'' zero differences obtain.

From this analysis of all the cross-category tetrads, we may conclude that depending upon the set-up, they furnish presumptive evidence of the *existence* of V or N, N or S, and V or S. This tentative conclusion must be modified by two considerations. First, even where Kelley's 16th Proposition is satisfied, the location of the

TABLE XVIII

1. Number Series
2. Arithmetic Reasoning
3. Computation
4. Otis
5. Beta
6. International*
7. Goodenough
8. Performance

a. Tetrads Arising from Two "Best" Tests of N and S

	Boys			*Girls*		
	t_{1245}	t_{1254}	t_{1452}	t_{1246}	t_{1264}	t_{1462}
td.	.2251	.2011	−.0240	.1386	.1247	−.0139
P.E.$_t$	±.0346	±.0371	±.0124	±.0293 ±.0297 ±.0308	±.0302 ±.0306 ±.0318	±.0034 ±.0034 ±.0035
$\frac{t}{P.E._t}$	6.5058	5.4205	1.9355	4.7304 4.6667 4.5000	4.1291 4.0752 3.9214	4.0882 4.0882 3.9714

b. Tetrads Arising from Two "Poorest" Tests of N and S

	Boys			*Girls*		
	t_{1378}	t_{1387}	t_{1783}	t_{1378}	t_{1387}	t_{1783}
td.	.0899	.1344	.0445	.0582	.0413	−.0169
P.E.$_t$	±.0336	±.0282	±.0129	±.0254	±.0275	±.0168
$\frac{t}{P.E._t}$	2.6756	4.7660	3.4496	2.2913	1.5018	1.0060

* P.E.'s for International based on three N's for the girls.

group factors cannot be ascertained. Second, when zero differences obtain in tetrads composed of two pairs of tests from two categories, an overlapping of the tentatively established central factors of each is indicated.

4. MEAN TETRADS

In order to investigate more fully the location, distinctiveness and magnitude of the C bond existing among the tests of any one category, the mean tetrads arising from the cross-category correlations were computed. This also gives additional indication of the effect produced on any one C by introducing into its tetrad set-ups tests of difference categories.

These average values were obtained in three ways, and the results are indicated in sections a, b, and c of the three tables which follow. An explanation of the methods is appended:

a. Each tetrad-difference was analyzed back into its component r's so that $t_{abcd} = .0620$ would be credited positively to ab and cd and negatively to ac and bd. For each r the td. values which express the strength of bond were summed algebraically and averaged. These averages are represented in the cells of the tables. Averages were then computed for all the tests of both categories: for the tests of each of the two categories, and for the correlations between a test of each category. The sigma of each of these averages was calculated as the best available method of evaluating the significance of the difference from zero.

b. The averages were also computed for the tetrad differences arising from two pairs of variables drawn from two categories simply as differences, without reassigning them to their constituent r's. These averages were computed in order to determine whether they would be of the type described by Kelley in his 16th Proposition, *i.e.,* two tetrads equal to each other and the third to zero, which would indicate in addition to the factor common to all the tests, a group factor common to either pair, or as in this case, a factor common to all the tests of one category.

c. The tetrad-values for all correlations between pairs of tests from *one* category in those tetrads which were composed of *two* pairs of tests of any *two* categories were then averaged and these averages analyzed in terms of their standard deviations. The significance of the ratio between the average of the tests of any one category and its S.D. was taken as an index of the differentiability of the bond uniting those tests. A separate analysis of this type of tetrad was necessary because in a tetrad composed of a single test of one category and three of another, the possible existence of differential factors might very well be obscured, whereas between pairs of tests from two categories they should appear at their maximum strength. The S.D. of each average was again taken as the best available measure for the determination of the significance of the average tetrad value from zero. This technique gave results which jibed nicely with the factors suggested by the result of the tetrad analysis of the "best" and "poorest" tests.

An analysis of the values in the three following tables yields the following findings:

A. *Verbal-Numerical*

When the tetrads arising from all seven tests are considered, the general averages of .0347 for the boys and .0265 for the girls indicate the presence of a bond underlying *both* the verbal and the numerical abilities. From this analysis the number tests are apparently more closely tied together than the verbal to judge by the higher averages which they yield.

TABLE XIX

1. Vocabulary
2. Analogies
3. Sentence Completion
4. Reading
5. Number Series
6. Arithmetic Reasoning
7. Computation

a. Mean Tetrads Arising from All Verbal and Numerical Tests
Above Diagonal: Boys, Below Diagonal: Girls

Test	*1*	*2*	*3*	*4*	*5*	*6*	*7*
1		.0066	.0061	.0996	−.0405	−.0062	−.0516
2	.0162		−.0130	.0211	.0239	−.0228	−.0057
3	.0331	−.0243		.0431	−.0156	−.0132	−.0088
4	.0340	.0138	.0612		−.1151	−.0289	−.0097
5	−.0570	.0827	.0262	−.0559		.0782	.0837
6	−.0378	.0319	−.0311	−.0440	.0815		−.0046
7	.0115	−.0455	−.0268	.0046	.0522	.0025	

	Boys		*Girls*	
		S.D.		*S.D.*
Average	.0129	.0866	.0061	.0414
Average V	.0273	.0366	.0223	.0260
Average N	.0524	.0404	.0454	.0326
Average VN	.0245	.0328	.0118	.0414

b. Mean Values for t_{abcd}, t_{abdc} and t_{acdb}

	Boys	*Girls*
t_{abcd}	.0719	.0737
t_{abdc}	.0996	.0750
t_{acdb}	.0277	.0013

c. Mean Values for Factor Bonds for r in Tetrads with Two Verbal and Two Number Tests

Variables	*Boys*	*Girls*
12	.0649	.0604
13	.0773	.0907
14	.1526	.0931
23	.0381	.0356
24	.0743	.0645
34	.1076	.1019
56	.1100	.1090
57	.1078	.0757
67	.0395	.0385

		S.D.		*S.D.*
Average V	.0858	.0362	.0744	.0229
Average N	.0858	.0326	.0744	.0288

$$\frac{\text{Av.V}}{\text{S.D.}} = 2.3702 \qquad \frac{\text{Av.V}}{\text{S.D.}} = 3.2489$$

$$\frac{\text{Av.N}}{\text{S.D.}} = 2.6319 \qquad \frac{\text{Av.N}}{\text{S.D.}} = 2.5833$$

TABLE XX

1. Vocabulary
2. Analogies
3. Sentence Completion
4. Reading
5. Otis
6. Beta
7. International
8. Goodenough
9. Performance

a. Mean Tetrads Arising from All Verbal and Spatial Tests
Above Diagonal: Boys, Below Diagonal: Girls

Tests	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
1		.0785	.0979	.1466	-.0702	-.1081	-.0290	-.0256	-.0925
2	.0421		.0643	.0744	-.0581	-.0724	-.0052	-.0571	-.0263
3	.0693	.0216		.1117	-.0594	-.0822	-.0413	-.0443	-.0493
4	.0719	.0254	.0789		-.0563	-.0856	-.0124	-.0707	-.1077
5	-.0758	-.0400	-.0661	-.0132		.0905	.0333	.0526	.0708
6	-.0697	-.0209	-.0636	-.0084	.0950		.0656	.0792	.1143
7	-.0170	.0292	-.0251	-.0736	.0532	.0538		-.0066	.0071
8	-.0272	-.0117	.0143	-.0424	.0067	-.0089	.0091		.0835
9	-.1813	-.0218	-.0054	.0380	.0469	.0231	-.0279	.0412	

	Boys		*Girls*	
		S.D.		*S.D.*
Average	-.0027	.0611	-.0022	.0148
Average V	.0952	.0276	.0515	.0229
Average S	.0590	.0360	.0292	.0337
Average VS	-.0577	.0286	-.0379	.0391

b. Mean Values for t_{abcd}, t_{abdc} and t_{acdb}

	Boys	*Girls*
t_{abcd}	.2049 (.2116)*	.1123 (.1321)
t_{abdc}	.1965 (.2002)	.0869 (.1005)
t_{acdb}	-.0079 (-.0107)	-.0247 (-.0316)

TABLE XX—(*Continued*)

c. Mean Values for Factor Bonds for r in Tetrads with Two Verbal and Two Spatial Tests

Variables	*Boys*		*Girls*	
12	.1840	(.1894)*	.0879	(.1041)
13	.2081	(.2134)	.1220	(.1467)
14	.2377	(.2407)	.1199	(.1386)
23	.1767	(.1869)	.0593	(.0704)
24	.1801	(.1830)	.0809	(.0893)
34	.2177	(.2220)	.1232	(.1412)
56	.2878		.2150	
57	.1986		.1664	
58	.1834		.0668	
59	.2056		.1196	
67	.2543		.1477	
68	.2187		.0430	
69	.2555		.0823	
78	.0927		.0514	
79	.1323		.0256	
89	.1784		.0741	
		S.D.		*S.D.*
	(.2059	.0212)	(.1151	.0292)
Average V	.2007	.0225	.0989	.0247
Average S	.2007	.0557	.0982	.0579

Boys:

$$\frac{\text{Av. V}}{\text{S.D.}} = \overset{(9.7123)}{8.9200}$$

$$\frac{\text{Av. S.}}{\text{S.D.}} = 3.6032$$

Girls:

$$\frac{\text{Av. V}}{\text{S.D.}} = \overset{(3.9418)}{4.0040}$$

$$\frac{\text{Av. S.}}{\text{S.D.}} = 1.7133$$

* All figures in parentheses indicate values when Performance is omitted.

TABLE XXI

1. Number Series
2. Arithmetic Reasoning
3. Computation
4. Otis
5. Beta
6. International
7. Goodenough
8. Performance

a. Mean Tetrads Arising from All Numerical and Spatial Tests
Above Diagonal: Boys, Below Diagonal: Girls

Tests	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
1		.1033	.0528	–.0352	–.0523	–.0035	–.0097	–.0563
2	.0181		.1008	–.0409	–.0882	.0278	–.0221	–.0663
3	.0409	.0503		–.0368	–.0183	.0167	–.0769	–.0259
4	–.0609	–.0522	–.0641		.0316	.0172	.0345	.0368
5	–.0151	–.0569	–.0368	.0702		.0212	.0405	.0661
6	–.0060	–.0147	–.0054	.0575	.0069		–.0303	–.0183
7	.0011	–.0013	–.0230	.0186	–.0283	.0076		.0632
8	.0184	–.0457	–.0423	.0318	–.0231	–.0585	.0281	

	Boys		*Girls*	
		S.D.		*S.D.*
Average	.0115	.0480	–.0066	.0344
Average N	.0856	.0232	.0364	.0135
Average S	.0263	.0294	.0111	.0374
Average NS	–.0325	.0317	–.0270	.0244

b. Mean Values for t_{abcd}, t_{abdc} and t_{acdb}

	Boys	*Girls*
t_{abcd}	.1314 (.1382)*	.0636 (.0821)
t_{abdc}	.1254 (.1184)	.0458 (.0729)
t_{acdb}	–.0160(–.0198)	–.0179(–.0093)

TABLE XXI—(*Continued*)

c. Mean Values for Factor Bonds for r in Tetrads with Two Number and Two Spatial Tests

Variables	*Boys*		*Girls*	
12	.1468	(.1437)*	.0417	(.0732)
13	.1001	(.0992)	.0563	(.0713)
23	.1385	(.1420)	.0661	(.0880)
45	.1772		.1534	
46	.1195		.1455	
47	.1298		.0758	
48	.1378		.0809	
56	.1424		.0475	
57	.1472		.0047	
58	.1741		–.0029	
67	.0536		.0381	
68	.0715		–.0353	
78	.1312		.0393	
		S.D.		*S.D.*
	(.1283	.0200)	(.0775	.0057)
Average N	.1285	.0201	.0547	.0100
Average S	.1284	.0376	.0547	.0645

$$\frac{\text{Av. N}}{\text{S.D.}} = \overset{(6.4150)}{6.3930} \qquad \frac{\text{Av. N}}{\text{S.D.}} = \overset{(13.5965)}{5.4700}$$

$$\frac{\text{Av. S}}{\text{S.D.}} = 3.4149 \qquad \frac{\text{Av. S}}{\text{S.D.}} = .7255$$

* All figures in parentheses indicate values when Performance is omitted.

When we consider the three average values for the tetrads arising from a pair of verbal and a pair of numerical tests, we find that while t_{abcd} and t_{abdc} are not very large, there is a tendency for them to equal each other and for t_{acdb} to be equal to zero, especially in the case of the girls.

When these tetrad values are reassigned to their constituent pairs of variables, an evaluation of the averages for the tests in each category in terms of the S.D. indicates a reliable bond only in the case of the verbal tests with the girls. The chances are sufficiently great to suggest the possibility of the existence of differential factors V and N, but not sufficiently great to be conclusive.

The following pairs of tests give especially large tetrad-values —for the boys, Vocabulary-Reading, Sentence Completion-Reading, Number Series-Arithmetic Reasoning, and Number Series-Computation; for the girls, Sentence Completion-Reading, Number Series-Arithmetic Reasoning, and to some extent, Vocabulary-Sentence Completion and Vocabulary-Reading.

B. *Verbal-Spatial*

The averages arising from all verbal and spatial tests are .0322 and .0143 respectively for the sexes. The averages for correlations between verbal and spatial tests are negative and in this set-up indicate to some extent a degree of unlikeness or opposition between the two abilities measured. Nevertheless there is something common to both the verbal and spatial tests, as indicated by the general averages. The average of the verbal test values for the boys is significant in terms of its S.D., .0952 S.D. .0276 = 3.4493. A special bond therefore exists among the verbal tests. The spatial tests give average tetrad-values lower than the verbal tests for both sexes.

The mean values for t_{abcd} and t_{abdc} arising from combinations of two verbal and two spatial tests are quite large especially for the boys and equal to each other, approximately, so that t_{acdb} tends to be about zero. The existence of either V or S is therefore indicated. These values increase somewhat, especially for the girls, when Performance is excluded.

When the tetrad-values for the verbal pairs and the spatial pairs are averaged and evaluated in terms of their S.D.'s, evidence is apparent for the differential existence of V and S for the boys, and of V for the girls. The V bond is somewhat stronger for the boys, when Performance is excluded.

The boys show strong linkages between all pairs of verbal tests and all pairs of spatial tests, except perhaps the Goodenough-Performance; the most strongly linked pairs for the girls are Vocabulary-Sentence Completion, Vocabulary-Reading and Sentence Completion-Reading, Otis-Beta, Otis-International, Otis-Performance and Beta-International.

C. *Numerical-Spatial*

The averages of the mean tetrads of the numerical and spatial tests, .0265 and .0068 again point to a common factor between these categories. In addition, the boys give evidence of a special number bond as the average value for the numerical tests is 3.6897 in terms of its S.D. Again the average value arising from one numerical and one spatial test is negative, indicating an opposition between these abilities.

t_{abcd} and t_{abdc} are higher for the boys than for the girls and tend to be more nearly equal to each other. Excluding Performance tends to give more of an increase in the average t_{abcd} and t_{abdc} for the girls than for the boys. The presence of N or S is more strongly suggested by the boys' results.

The averages for the numerical pairs give evidence of a differential N for both boys and girls and this evidence is stronger, especially for the girls, when Performance is excluded. The differential existence of S is indicated for the boys only, in the case of the girls its independence is negligible.

The boys give large values for all pairs of numerical tests and all pairs of spatial tests except Goodenough-International and Performance-International; the girls yield insignificant results for all numerical test pairs and all spatial test pairs except Otis-Beta and Otis-International.

5. THE RESULTS OF THE SEVERAL METHODS OF TETRAD ANALYSIS

The results of the cross-tetrad analyses are summarized schematically in Table XXII: + indicating the presence of a differential factor and – its absence, as far as the present analysis goes.

TABLE XXII

THE EXISTENCE AND LOCATION OF DIFFERENTIAL FACTORS SCHEMATICALLY REPRESENTED

a. Single Tetrads Composed of Pairs of "Bests" and "Poorests"

Categories	*Tests*	*Boys*		*Girls*	
		t_{abcd}	t_{abdc}	t_{abcd}	t_{abdc}
Verbal-Number	Best	+	–	+	+
	Poorest	–	+	+	–
Verbal-Spatial	Best	+	+	+	+
	Poorest	+	+	–	–
Number-Spatial	Best	+	+	+	+
	Poorest	–	+	–	–

b. Tetrad-Averages for All Tests of a Single Category

	Boys			*Girls*		
	Verbal-Number	*Verbal-Spatial*	*Number-Spatial*	*Verbal-Number*	*Verbal-Spatial*	*Number-Spatial*
V	–	+		+	+	
N	–		+	–		+
S		+	+		–	–

In Table XXIIa for the boys, evidence for the existence of either factor V or S is strongest in the verbal-spatial tetrads, while N or

S is next most strongly indicated. The results for the verbal-number analysis are ambiguous. For the girls, the most significant differences are found for the verbal-number tests. The presence of either V or S or N or S is indicated for the best test-pairs, while among the poorest number-spatial and verbal-spatial tests overlapping of the C's occurs.

In Table XXIIb the boys give evidence both times of a spatial factor and the girls of a verbal factor. N is not differentiated for either sex in the verbal-number tests and for both sexes in the number-spatial tests. For the boys V is indicated in the verbal-spatial tetrads, for the girls S is not differentiable in either set-up.[16]

To compare these results with those obtained for averages and correlations, it may be noted that the existence of a special link or factor among tests of a similar kind does not make for reliable superiority in that factor. The boys were superior, although not reliably so, in two of the language tests, Analogies and Vocabulary; yet the girls give more reliable indications of the presence of a verbal factor. The boys give equally reliable evidence for the existence of S and yet, although they were superior in three of the spatial tests, Otis, Beta and International, they were reliably superior in only two, Otis and Beta.

These sex differences which are extremely interesting may well be due to the fact that boys are early encouraged and accustomed to the use of tools and construction materials and activities and this training in the manipulation of spatial relationships is reinforced by participation in games which also involve more elaborate spatial formations. Girls, on the other hand, are more likely to be provided with reading materials for recreational purposes.

The lack of conclusive evidence of the split-off of N as a factor discrete and independent from V is probably to be attributed, as were the relatively large verbal-number intercorrelations, to the fact that with these young children number processes are still to a large extent dependent upon specifically taught verbalizations. That this split-off between V and N may eventually occur is an assumption rendered plausible by the results for older subjects in a very similar sampling of the population, *cf.* Schneck (48).

The general averages of all tetrad-differences arising from tests of any two categories (Sections a of Tables XIX to XXI) must now be considered. Since these averages approximate zero, they indicate in themselves, the community of a single undifferentiable factor linking the measures of two apparently different types of abilities. The further necessary inference from these results that the ability measured by the tests of one category is in some degree, at least, the same as that measured by the tests of another category must be qualified by the following consideration: The general average is determined by tetrads composed of a single variable from one category and three of another. A differentiable factor existing between the two tests of the same category would tend to be balanced by the negative relationship which is so frequently apparent in Section a of Tables XIX to XXI between that pair of variables composed of a single test from each category. The averages of such pairs of variables are negative in the NS and VS pairs for both boys and girls. Even in these set-ups, exceptionally strong bonds will persist, *e.g.,* the differentiable V and N bonds for the boys in Tables XX and XXI respectively. If these statistical considerations are disregarded, however, and the existence of a unitary function among the tests of any two categories considered as established, the nature of this unitary function remains to be determined.

---

[16] Similar results are reported by Spearman (51, p. 229) concerning the absence of a spatial group factor with girls.

The alpha factor of heterogeneity in sex, nurture and maturity postulated by Kelley as a unifying principle does not apply here inasmuch as these variables were experimentally controlled. Whatever it is that is common to any two of our categories cannot very well be a product of heterogeneity.

Spearman has advanced two interpretations of the zero-tetrad relationship. The earlier of these (51, p. 75) claims only the establishment of a general factor or *g,* not primarily ''as any concrete thing but only a value of magnitude. Further, that which the magnitude measures has not been defined by declaring what it is like, but only by pointing out where it can be found. It consists of *just that constituent whatever it may be that is common to all the abilities intercorrelated by the tetrad equation*—a defining of *g* by site rather than by nature.'' (Italics ours.) Nevertheless, this magnitude is measured, albeit roughly, by current tests of intelligence, and is identifiable, physiologically, with mental energy. Later, however, *g* is definitively identified with ''general intelligence'' or ''general ability.'' (53)

In the interests of a parsimonious interpretation and mindful that the factors indicated by certain relationships among the tetrad constituents are not necessarily the factors which *per se* cause the relationship, (*cf.* Kelley, 34), we offer, in accordance with Spearman's earlier claims, the following considerations as accounting for the community of functions revealed by our zero average tetrads:

1. The influences tending to make for high intercorrelations (as discussed in Chapter III) *e.g.,* the similarity of the general testing procedures and situations, the use of pencil and paper, the presence of the same examiner for some of the tests, the similarity of timing and test length, the prevalence of the same room conditions, the presentation of the same oral instructions, and very similar motivations would also tend to make for zero tetrads (*cf.* Schneck, 48, p. 36 and Peatman, 42, p. 48). These conditions give relatively less weight to the group functions. It should be emphasized that although these similarities would also account for the existence of a single factor among the tests of any one category, they would tend to vitiate rather than to produce the differentiable factors found here.

2. In tetrads composed of three tests from one category and a fourth from a second category (*e.g.,* VVVN) zero averages may have been produced when a large positive tetrad bond between the two tests of the same category were offset by the relationship between the two tests of two different categories.[17] Especially is this true in set-ups where these latter bonds may have been caused by the fact that the same mental processes were involved. Examinations of Table II will show, for example, that there were verbal and spatial analogies tests, verbal and spatial directions tests, that some computation and arithmetical reasoning abilities were involved in the cube-counting of International, while completion was involved in the Number Series, the Prediction of Outcomes sub-test in the Reading test and in the spatial narrative-sequence sub-tests. The significant tetrad-bond possibly existing between Reading and Computation for the boys has already been discussed (p. 43). The inadequacy of zero means tetrads to prove the absence of group factors has been pointed out by Findley (13, App. D). Brigham analyzes his three tests of verbality and one of number and concludes that the insignificant zero tetrads which they yield do not indicate that the verbal and number items are "measuring the same function to a degree such that the two scores could be added to a single index without losing valuable information" (4, p. 374).

3. Symbols were the common medium of exchange in both the test-administration and test-composition[18]—these were verbal in the

---

[17] *Cf.* Brigham, "All tetrads have to be small if they arise from combinations of three highly homogeneous tests and one extraneous one." 4, p. 374.

[18] Except to a certain extent in the Performance.

former and verbal, numerical or spatial in the latter. This prevalence of symbolism may well have acted as a unifying factor, so that the constituent common to all the variables of any two categories may be some factor of the nature "manipulation of symbolic relations." Recent evidence (*cf.* Anastasi, 2, and Smith, 49) suggests that when these symbols were of the same character, *i.e.*, verbal, numerical or spatial, group factors resulted among the similar materials.

These qualifications are not to be construed as denials of the existence of g as defined by Spearman. They are rather an attempt to account for any common functions which may exist between tests of two categories in the light of observable and observed test conditions. Similar procedures and constructions seem to be a more desirable explanation than the projection of statistical findings into hypothetical psychological entities. The discrepant nature of some of our own results indicates that to date the factors found in any research are at least as much a product of those forces as they are of the actual mental organization. We may cite the vanishing of tetrads in certain set-ups and of group factors in others, and the difference between the present results and those obtained by other experimenters when different types of tests, subjects, experimental conditions and controls and statistical techniques were employed.

Certain implications for the principles and practises of mental test construction and use are suggested by our results (*cf.* 18, p. 235–236). That factor which is most concerned with the particular activity to be diagnosed or predicted should be determined. Batteries should then be constructed in which sub-tests of this factor would be weighted according to their correlation with it. If more than one factor has a significant correlation with the activity in question, the sub-test measures of the several factors should be weighted according to their contributions to it. In this way, whether the test scores be used in the scholastic classification of children, the treatment of behavior difficulties, nature-nurture investigations or in racial comparisons, greater allowance will be made for the rôle played by the relevant, identifiable, and independent entities in the organization of mental abilities. There is no tacit assumption that a single unitary function and nothing else is being equitably measured by a system of sub-tests, or by a hodge podge of tests grouped into one or more batteries.

CHAPTER VI

SUMMARY AND IMPLICATIONS

1. A group of 189 boys and 206 girls, homogeneous with respect to age, parentage, schooling and socio-economic environment, were tested with a battery of four verbal, three numerical and five spatial-manipulation tests, the results for the sexes being treated separately.

2. Reliable sex differences in ability were found in Arithmetic Reasoning, Otis and Army Beta in favor of the boys. In Goodenough Drawing-a-Man test the girls were superior. Sex differences in variability were negligible.

3. The average correlation of all 12 tests was the same for both sexes, viz., .5580. The verbal tests correlated highest among themselves, the numerical tests were intermediate, and the tests of spatial ability showed the lowest correlation *inter se*. The verbal and numerical tests correlated higher among themselves than the tests of any other two categories.

4. Within a given category, *i.e.*, verbal, numerical, or spatial-manipulation, tetrad or triad analysis indicated the presence of a central factor. These were denominated V, N, and S, respectively. The existence of V and N was substantiated by the boundary condition technique; but this technique did not establish S conclusively.

5. When the cross-tetrads resulting from two pairs of the *best*[19] tests of each category, or two pairs of the *poorest*[23] tests of each category were analyzed, the boys gave evidence of the existence of V or S in the verbal-spatial set-ups, and partial evidence of N or S in the number-spatial tetrads. The results for the verbal-number tetrads were ambiguous. The girls gave partial evidence for V or N in the verbal-number tetrads. In the combination of ''bests'' in the verbal-spatial and the numerical-spatial tetrads, results for the girls indicated V or S, or N or S respectively. The combinations of ''poorests'' for these categories gave zero differences.

6. The tetrad-difference averages which indicate the bonds between the tests of any two categories were all of zero order. This result demonstrates the possibility of a single common function throughout all such tests. The available explanations of this com-

---

[19] ''Best'' tests refers to those tests having highest correlations with their common factor. ''Poorest'' tests are those tests having lowest correlations with their common factor.

mon function, viz., Kelley's "alpha" factor of heterogeneity and Spearman's *g*-factor were discussed. The rôles played by background and motivation of subjects, experimental and statistical procedures, prevalence of symbolic materials, in the research set-up were also considered in the interpretation of the function common to any two test-categories.

7. Negative tetrad-bonds were observed for correlations between tests from two different categories.

8. When only those tetrads which were composed of pairs of tests from two categories were considered, the boys presented evidence of a differentiable spatial factor throughout the verbal-spatial and numerical-spatial tests. The girls exhibited a differentiable verbal factor in the verbal-number and verbal-spatial tests. A spatial factor was not differentiated for the girls in either the verbal-spatial or the numerical-spatial tetrads. Both sexes gave a differentiable number factor in the number-spatial set-ups, but not in the verbal-number set-ups. The boys' results indicated V as a group factor in the verbal-spatial tests but not throughout the verbal-number tests. A tentative explanation of the differentiable group factors thus located was advanced on the basis of similarity of test *materials* of any one category.

Since Spearman's pioneer work in 1904, research in the field of mental organization has increased apace. The problem of the constitution of intelligence has, however, been the center of interest and controversy. The mental abilities tested have increased in number and variety and the subjects employed have been characterized by greater diversity in maturity, heredity, and general nurture. At the same time, the statistical assumptions and methodology have undergone continuous analysis and development.

On the basis of the present study, both in itself and in relation to previous research, several interpretations may be advanced to account for existing discrepancies. In the first place, the identity of the function tested is not an absolute entity. The fact that two tests present the same type of task to the subject does not mean that those tests measure the identical ability. A vocabulary test devised for 9-year-old children by one experimenter does not necessarily tap the same ability as a vocabulary test devised for 12-year-old children. To be fairly certain that the same ability was being measured in both groups, a single test must be devised adequate to test both the dullest 9-year-old and the brightest 12-year-old. Correlations and pattern analyses would then have a more consistent meaning from group to group.

Secondly, as concerns the hereditary and environmental constitution of the subjects employed, methods of differentiating the effects of these two forces are still in a preliminary experimental stage. Since neither force functions in a medium unadulterated by the other, the problem of interpreting results obtained for any two different groups of subjects becomes the more complex.

Moreover, age plays a rôle which still further complicates the situation. On the heredity side it may well be that mental processes undergo maturational transitions from generalized capacities for intellectual response to selectively consolidated functional unities—verbal, numerical, or spatial in nature. Such a process would be analogous to the maturation of motor processes and activities from the diffused, uncoordinated and generalized to the precise, coordinated, and specialized. Environmentally, and with at least as much certainty, such concomitants of age as increasing selection through the educational program, and the shifts in social pressures —notably, as they differentially affect the sexes, also exert an immeasurable but vital influence in such results as have been reported.

Additional problems such as race, and others more difficult to dichotomize as hereditary or as environmental, such as differentiation of interests, socio-economic status, etc., also confront us. Thus it becomes difficult to fit the results so far obtained into any consistent reflection of the organization of mental abilities. Then too, since such results are at least as much again a product of the statistical techniques which yield them as they are of the foregoing considerations, then to the extent that these techniques are discrepant or fallible, consistency and ultimate insight must await their unification and perfection.

Finally, the vexing question of the effects of heterogeneity must be taken into account. Not only between groups, but within a group differences in age, socio-economic status and parental backgrounds and traditions complicate indeterminably both the resolution of mental abilities into accurately representative factor patterns and the analyses of the cause, nature, and significance of such differentiation as has been found.

Despite these difficulties, more than a beginning has been made. The emphasis is no longer on innumerable specific abilities nor on a single all-pervasive one, but rather on linkages among the former sufficiently powerful and coherent to detract from the importance of the latter.

BIBLIOGRAPHY

1. Anastasi, A. A Group Factor in Immediate Memory. *Arch. Psychol.*, 1930, 120.
2. Anastasi, A. Further Studies on the Memory Factor. *Arch. Psychol.*, 1932, 142.
3. Bernstein, E. Quickness and Intelligence. *Brit. J. Psychol. Monog.*, 1924, No. 7.
4. Brigham, C. C. A Study of Error. New York, College Entrance Examination Board, 1932, p. 384.
5. Brigham, C. C. Validity of Tests in Examination of Immigrants. *Indust. Psychol.*, 1926, 1, 413–417.
6. Brigham, C. C. Intelligence Tests of Immigrant Groups. *Psychol. Rev.*, 1930, 37, 158–165.
7. Burt, C. Experimental Tests of General Intelligence. *Brit. J. Psychol.*, 1909, 3, 94–177.
8. Colvin, S. S., and Allen, R. D. Mental Tests and Linguistic Ability. *J. Educ. Psychol.*, 1923, 14, 1–20.
9. Cureton, E. E., and Dunlap, J. W. Some Effects of Heterogeneity on the Theory of Factors. *Am. J. Psychol.*, 1930, 42, 608–620.
10. Davey, C. M. A Comparison of Group Verbal and Pictorial Tests of Intelligence. *Brit. J. Psychol.*, 1926, 17, 27–49.
11. Dodd, S. C. International Group Intelligence Tests. Princeton, N. J., 1926.
12. Dunlap, J. W., and Kurtz, A. K. Handbook of Statistical Nomographs, Tables, and Formulas. Yonkers, N. Y., World Book Co., 1932, p. 163.
13. Findley, W. G. Specialization of Verbal Facility at the College Entrance Level. N. Y., Teachers College, 1933, p. 76.
14. Fisher, R. A. Statistical Methods for Research Workers. London, Oliver, 1930, p. 284.
15. Flugel, J. C. Practise, Fatigue, and Oscillation. *Brit. J. Psychol., Monog. Suppl.*, 1928, 4, No. 13, p. 92.
16. Garnett, J. C. M. General Ability, Cleverness and Purpose. *Brit. J. Psychol.*, 1919, 9, 345–366.
17. Garrett, H. E. Statistics in Psychology and Education. N. Y., Long, 1929, p. 317.
18. Garrett, H. E., and Anastasi, A. The Tetrad Difference Criterion and the Measurement of Mental Traits. *Ann. N. Y. Acad. Sc.*, 1932, 33, 233–282.
19. Gates, A. I. Correlations of Achievement in School Subjects with Intelligence Tests and Other Variables. *J. Educ. Psychol.*, 1922, 13, 129–139, 277–285.
20. Gates, A. I. A Series of Tests for the Measurements and Diagnosis of Reading Ability in Grades 3–8. *Teach. Coll. Rec.*, 1926, 28, No. 1.
21. Gates, A. I. Methods of Constructing and Validating the Gates Reading Tests. *Teach. Coll. Rec.*, 1927, 29, 148–159.
22. Gaw, F. A Study of Performance Tests. *Brit. J. Psychol.*, 1924, 15, 375–392.
23. Gaw, F. Performance Tests of Intelligence. I.F.R.B. Report, 1925, No. 31.
24. Goodenough, F. L. Measurement of Intelligence by Drawing. Yonkers, N. Y., World Book Co., 1926, p. 177.
25. Goodenough, F. L. A New Approach to the Measurement of the Intelligence of Young Children. *Ped. Sem. and J. Genet. Psychol.*, 1926, 33, 185–211.
26. Halpern, F. C. The Intelligence of Bilingual Jewish Girls as Measured by the Stanford Binet and the Pintner-Paterson Performance Tests. M.A. Thesis, Columbia University, 1933. (Unpublished.)
27. Hart, B., and Spearman, C. General Ability, its Existence and Nature. *Brit. J. Psychol.*, 1912, 5, 51–84.
28. Herring, J. P. Verbal and Abstract Elements in Intelligence Examinations. *J. Educ. Psychol.*, 1921, 12, 511–517.

29. Holzinger, K. J. Statistical Resume of the Spearman Two Factor Theory. Univ. Chicago Press, 1930, p. 43.
30. Holzinger, K. J. On Factor Theory. Conference on Individual Differences in Special and General Abilities. Wash., D. C., Nat. Res. Council, 1931. (Unpublished.)
31. Johnson, B., and Schriefer, L. A. Comparison of Mental Age Scores obtained by Performance Tests and Stanford Revision of the Binet-Simon Scale. *J. Educ. Psychol.*, 1922, 13, 408–417.
32. Kefauver, G. N. Need of Equating Intelligence Test Quotients Obtained from Group Tests. *J. Educ. Res.*, 1929, 19, 92–101.
33. Kelley, T. L. Crossroads in the Mind of Man. Stanford, Calif., Stanford Univ. Press, 1928, p. 238.
34. Kelley, T. L. *In* Conference on Individual Differences. Washington, D. C., Nat. Res. Council, 1931. (Unpublished.)
35. Klineberg, N. V. Bilingualism and Intelligence in 10-year-old Italian Girls, M. A. Thesis, Columbia University, 1932. (Unpublished.)
36. Lankes, W. Perseveration. *Brit. J. Psychol.*, 1915, 7, 387–419.
37. Meili, R. Recherches sur les Formes d'intelligence. *Arch. de Psychol.*, 1930, 22, 201–284.
38. Morganthau, D. R. Some Well-Known Mental Tests Evaluated and Compared. *Arch. Psychol.*, 1922, No. 52.
39. Murdock, K., Maddow, D., and Berg, A. A Study of Relation between Intelligence and the Acquisition of English. 27th Yearbook Nat. Soc. Stud. Educ., 1928, Part 1, 342–353.
40. Otis, S. A. An Absolute Point Scale for the Measurement of Intelligence. *J. Educ. Psychol.*, 1918, 9, 239–261; 333–348.
41. Paterson, D. G., Elliott, R. M., Anderson, L. D., Toops, H. A., Heidbreder, E. Minnesota Mechanical Ability Tests. Minneapolis, U. of Minn. Press, 1930, p. 586.
42. Peatman, J. G. Factors Measured by the Thorndike Intelligence Examination for High School Graduates. *Arch. Psychol.*, 1932, No. 128.
43. Pintner, R., and Paterson, D. G. A Scale of Performance Tests. N. Y., Appleton, 1917, p. 218.
44. Pintner, R. Intelligence Testing. Rev. Ed., N. Y., Holt, 1931, p. 555.
45. Price, M. K. Reliability of International Mental Test. M.A. Thesis, Columbia University, 1932. (Unpublished.)
46. Root, W. T. Correlations between Binet Tests and Group Tests. *J. Educ. Psychol.*, 1922, 13, 286–293.
47. Schieffelin, B., and Schwesinger, G. C. Mental Tests and Heredity. N. Y., Galton Publ. Co., 1930, p. 298.
48. Schneck, M. M. R. The Measurement of Verbal and Numerical Abilities. *Arch. Psychol.*, 1929, No. 107.
49. Smith, G. M. Group Factors in Mental Tests Similar in Structure or in Material. *Arch. Psychol.*, 1933, No. 156.
50. Spearman, C. ''General Intelligence'' Objectively Measured and Determined. *Am. J. Psychol.*, 1904, 15, 201–293.
51. Spearman, C. The Abilities of Man. N. Y., Macmillan, 1927, p. 145.
52. Spearman, C. Disturbers of Tetrad Differences. Scales. *J. Educ. Psychol.*, 1930, 21, 559–573.
53. Spearman, C. In Psychologies of 1930. ed. Murchison, Worcester, Mass., Clark Univ. Press, 1930, p. 497.
54. Stead, H. G. Factors in Mental and Scholastic Ability. *Brit. J. Psychol.*, 16, 1926, 199–221.
55. Stephenson, W. Tetrad-Differences for Non-Verbal Subtests. *J. Educ. Psychol.*, 1931, 22, 167–185.
56. Stephenson, W. Tetrad-Differences for Verbal Sub-Tests. *J. Educ. Psychol.*, 1931, 22, 255–267.
57. Stephenson, W. Tetrad-Differences for Verbal Subtests Relative to Non-Verbal Subtests. *J. Educ. Phychol.*, 1931, 22, 334–350.
58. Stockton, J. L. The Definition of Intelligence in Relation to Modern Methods of Mental Measurement. *Psychol. Mono.*, 1921, 30, No. 137, p. 118.

59. Sylvester, R. H. The Form Board Test. *Psychol. Mono.*, 1913, 15, No. 65.

60. Terman, L. M., and Chamberlain, M. B. 23 Serial Tests of Intelligence and Their Intercorrelations. *J. Appl. Psychol.*, 1918, 2, 341–354.

61. Thompson, J. R. Boundary Conditions for Correlation Coefficients between Three and Four Variables. *Brit. J. Psychol.*, 1928, 19, 77–94.

62. Thorndike, E. L. Educational Psychology. Vol. 3, N. Y., Teach. Coll., 1913, p. 408.

63. Thorndike, E. L. On the Organization of Intellect. *Psychol. Rev.*, 1921, 28, 141–151.

64. Thorndike, E. L., and others. The Measurement of Intelligence. N. Y., Teachers College, 1926, p. 616.

65. Tryon, R. C. Multiple Factors *vs.* Two Factors as Determiners of Abilities. *Psychol. Rev.*, 1932, 39, 324–251.

66. Tryon, R. C. So-Called Group Factors as Determiners of Abilities. *Psychol. Rev.*, 1932, 39, 403–439.

67. Wallin, J. E. W. The Consistency Shown by Intelligence Ratings Based on Standard Tests and Teachers' Estimates. *J. Educ. Psychol.*, 1923, 14, 231–245.

68. Walters, F. C. Language Handicap and the Stanford-Binet. *J. Educ. Psychol.*, 15, 1924, 276–284.

69. Webb, E. Character and Intelligence. *Brit. J. Psychol.*, Monog. Suppl., 1915, 1, (3) p. 99.

70. Wilson, J. H. On the Nature of Intelligence. *J. Educ. Psychol.*, 1931, 22, 20–34.

71. Winsor, A. L. The Relative Variability of Boys and Girls. *J. Educ. Psychol.*, 1927, 18, 327–336.

72. Worthington, M. R. A Study of Some Commonly Used Performance Tests. *J. Appl. Psychol.*, 1926, 10, 216–227.

73. Yoakum, C. S., and Yerkes, R. M. Army Mental Tests. N. Y., H. Holt, 1929, p. 303.

APPENDICES

APPENDIX I

QUESTIONNAIRE

Name .... Date .... Class ....
Boy or Girl .... Date of Birth ....
Where were you born? .... Are you Jewish? ....
Where was your father born? ....
How long has your father been in the United States? ....
Where was your mother born? ....
How long has your mother been in the United States? ....
What is your father's work? ....
Does he own his own business? ....
What is your mother's work? ....
Did your father go to public school? .... High School? .... College? ....
Did your mother go to public school? .... High School? .... College? ....
Are you the oldest child? .... Second Child? .... Third Child? ....
Fourth Child? .... Fifth Child? .... Sixth Child? ....
What languages are spoken in your home? ....
Do you speak Jewish? .... Do you understand Jewish? ....
What language do your parents speak to you *most* of the time? ....
What other language do they speak to you besides that? ....
What language do your parents speak to your older brothers and sisters *most* of the time? ....
What language do your parents speak with each other *most* of the time? ....
What language do your uncles and aunts speak to you *most* of the time? ....
What language do your grandparents speak to you *most* of the time? ....
What language do you speak with your brothers and sisters? ....

APPENDIX II[a]

TESTING SCHEDULE AND EXAMINERS

	1–25–32 Monday		*1–26–32 Tuesday*		*1–27–32 Wednesday*		*1–28–32 Thursday*		*1–29–32 Friday*		*2–1–32 Monday*		*2–2–32 Tuesday*		*2–3–32 Wednesday*	
Class	*A.M.*[b]	*P.M.*	*A.M.*	*P.M.*	*A.M.*	*P.M.*	*A.M.*	*P.M.*	*A.M.*	*P.M.*	*A.M.*	*P.M.*	*A.M.*	*P.M.*	*A.M.*	*P.M.*
3A^1	No. Series[c] Vocab.		Comput.		Otis (L.A.)	Begin Beta (N.H.)	Finish Beta		Arith. Reas.		Analogies		Sent. Compl. Int.II[d] (L.A.)	Int.II		Int.I (N.H.)
3A^2	No. Series Vocab.		Comput.	Otis Beta (N.H.)		Int.II (N.H.)	Read.		Begin Int.I (N.H.) Arith. Reas.	Finish Int.I (N.H.)	Analogies		Sent. Compl.			
3A^3	No. Series Vocab.		Comput.	Otis (L.A.)		Beta (N.H.)	Read.		Arith. Reas.		Analogies Begin Int.II (N.H.)	Finish Int.II	Sent. Compl. Int.I (N.H.)	Int.I		
3A^4	Otis (N.H.) No. Series Vocab.		Comput. Beta (N.H.)			Int.II (N.H.)	Read.		Arith. Reas. Int.I (N.H.)		Analogies Finish Int.I		Sent. Compl.			
3B^1	No. Series Vocab.		Comput.		Otis (L.A.)		Read.	Beta (L.A.)	Arith. Reas.		Analogies Begin Int.II (L.A.)	Finish Int.II	Sent. Compl.		Int.I (N.H.)	Int.I
3B^2	No. Series Vocab.		Comput.	Otis (L.A.)	Beta (N.H.)		Read.		Begin Int.II (L.A.) Arith. Reas.	Finish Int.II	Analogies	Begin Int.I (N.H.)	Finish Int.I Sent. Compl.			
3B^3	No. Series Vocab.		Comput.		Begin Otis (L.A.)	Finish Otis	Read.	Beta (L.A.)	Arith. Reas.		Analogies Begin Int.II (L.A.)	Finish Int.II	Sent. Compl.	Begin Int.I (L.A.)	Finish Int.I (N.H.)	

APPENDIX II[a]—(*Continued*)

TESTING SCHEDULE AND EXAMINERS

	1–25–32		*1–26–32*		*1–27–32*		*1–28–32*		*1–29–32*		*2–1–32*		*2–2–32*		*2–3–32*	
	Monday		*Tuesday*		*Wednesday*		*Thursday*		*Friday*		*Monday*		*Tuesday*		*Wednesday*	
Class	*A.M.*[b]	*P.M.*	*A.M.*	*P.M.*	*A.M.*	*P.M.*	*A.M.*	*P.M.*	*A.M.*	*P.M.*	*A.M.*	*P.M.*	*A.M.*	*P.M.*	*A.M.*	*P.M.*
4A¹	No. Series Vocab.		Comput.			Otis (L.A.)	Read.		Arith. Reas.	Beta (L.A.)	Analogies	Begin Int.II (N.H.)	Sent. Compl. Finish Int.II	Begin Int.I (N.H.)	Finish Int.I	Finish Int.I
4A²	No. Series Vocab.		Comput.	Otis (N.H.)	Beta (N.H.)		Read.	Begin Int.II (N.H.)	Arith. Int.II	Finish Int.II	Analogies		Sent. Compl. Begin Int.I (N.H.)	Finish Int.I		
4A³	No. Series Vocab.		Comput.			Otis (L.A.)	Beta (L.A.) Read.		Arith. Reas. Begin Int.II (L.A.)	Finish Int.II	Analogies Int.I (N.H.)	Int.I	Sent. Compl.			
4B¹	No. Series Vocab.		Comput.			Otis (L.A.)	Read.	Beta (N.H.)	Arith. Reas.		Analogies	Int.II (L.A.) Int.II	Sent. Compl.		Int.I (L.A.) Int.I	
4B²	Otis (N.H.) No. Series Vocab.	Beta (N.H.)	Comput. Begin Int.II (N.H.)	Finish Int.II			Read.		Arith. Reas.		Analogies		Sent. Compl.			
4B³	No. Series Vocab.		Comput.			Otis (L.A.)	Read.		Arith. Reas.	Beta (L.A.)	Analogies		Begin Int.II (L.A.) Sent. Compl.	Finish Int.II	Int.I (L.A.)	Int.I

Note: Thursday and Friday, 2–4–32 and 2–5–32, questionnaires were filled out and absentees tested.

[a] Goodenough given previous Friday 1–22–32.

[b] A.M. = 9: 50–12: 10, P.M. = 1: 10–3: 50.

[c] Verbal and number tests administered by teachers.

[d] The demonstration charts for Book I of International were delayed in reaching the school. International I was accordingly administered first rather than to risk a shortage of time towards the end of the testing-program.

APPENDIX III

FREQUENCY DISTRIBUTIONS

1. Age

Scores	*Boys*	*Girls*
570–579	2	
560–569	3	1
550–559	2	4
540–549	3	2
530–539	5	3
520–529	9	7
510–519	9	14
500–509	21	17
490–499	13	13
480–489	26	23
470–479	16	18
460–469	12	21
450–459	14	18
440–449	16	19
430–439	16	20
420–429	7	8
410–419	8	7
400–409	7	9
390–399		1
380–389		1

2. Number Series

Scores	*Boys*	*Girls*
35–39	6	
30–34	13	13
25–29	33	21
20–24	40	53
15–19	39	54
10–14	30	30
5–9	17	26
0–4	11	9

3. Arithmetic Reasoning

Scores	*Boys*	*Girls*
55–59	1	1
50–54	25	9
45–49	48	27
40–44	47	46
35–39	26	45
30–34	19	42
25–29	15	20
20–24	4	10
15–19	3	4
10–14	1	2

4. Computation

150–159	3	2
140–149	9	17
130–139	22	20
120–129	23	25
110–119	18	19
100–109	31	27
90–99	22	29
80–89	17	22
70–79	18	16
60–69	12	15
50–59	8	8
40–49	3	4
30–39	3	2

5. Vocabulary

65–69	1	
60–64	1	2
55–59	11	6
50–54	25	22
45–49	31	37
40–44	32	43
35–39	37	40
30–34	14	24
25–29	13	11
20–24	5	9
15–19	12	6
10–14	4	0
5–9	3	3
0–4		3

6. Analogies

60–64	3	2
55–59	14	11
50–54	12	14
45–49	12	15
40–44	17	18
35–39	25	19
30–34	19	23
25–29	35	45
20–24	26	35
15–19	20	16
10–14	5	4
5–9	1	3
0–4		1

7. Sentence Completion

60–64	1	1
55–59	7	9
50–54	11	10
45–49	28	28
40–44	18	29
35–39	29	32
30–34	29	33
25–29	25	26
20–24	21	19
15–19	8	13
10–14	7	5
5–9	2	1
0–4	3	

8. Reading

95–104	2	4
85–94	10	7
75–84	8	13
65–74	13	9
55–64	25	24
45–54	27	35
35–44	36	44
75–34	28	33
15–24	25	29
5–14	15	8

APPENDIX III.—(*Continued*)

FREQUENCY DISTRIBUTIONS[e]

9. Otis

Scores	*Boys*	*Girls*
50–54	1	2
45–49	38	24
40–44	47	50
35–39	53	49
30–34	32	40
25–29	13	29
20–24	3	8
15–19	2	3
10–14		0
5–9		1

10. Beta

Scores	*Boys*	*Girls*
58.5–63.4	1	
53.5–58.4	1	1
48.5–53.4	6	3
43.5–48.4	19	9
38.5–43.4	31	17
33.5–38.4	41	36
28.5–33.4	43	53
23.5–28.4	28	52
18.5–23.4	14	29
13.5–18.4	4	4
8.5–13.4	1	2

11. International[f]

Scores	*Boys*	*Girls*
290–304	1	4
275–289	4	2
260–274	12	4
245–259	11	6
230–244	10	13
215–229	10	20
200–214	22	21
185–199	24	15
170–184	14	24
155–169	14	27
140–154	20	16
125–139	9	15
110–124	7	14
95–109	3	4
80–94		1

12. Goodenough

45–49		1
40–44	2	2
35–39	1	7
30–34	13	15
25–29	25	45
20–24	67	61
15–19	53	63
10–14	24	12
5–9	4	

13. Performance

550.0–569.9		1
530.0–549.9		
510.0–529.9	1	2
490.0–509.9	2	1
470.0–489.9	3	3
450.0–469.9	6	8
430.0–449.9	10	11
410.0–429.9	11	12
390.0–409.9	15	14
370.0–389.9	15	29
350.0–369.9	23	30
330.0–349.9	20	34
310.0–329.9	24	18
290.0–309.9	21	19
270.0–289.9	22	14
250.0–269.9	10	7
230.0–249.9	6	3

[e] Boys' N = 189, Girls' N = 206.

[f] Boys' N = 161, Girls' N = 186.

APPENDIX IV

PERFORMANCE SUB-TEST INTERCORRELATIONS

a. Boys

N = 189

	Healy A Time *1*	*Healy A Moves* *2*	*5 Figure Time* *3*	*5 Figure Errors* *4*	*Casuist Time* *5*	*Casuist Errors* *6*	*Knox Cube* *7*
Sigma ...	105.0780	20.7782	82.0375	12.5336	79.5445	12.1566	1.7627
Mean	169.4603	36.3757	133.8836	13.4233	138.0899	13.5291	6.1746
	1.	.8728	.3053	.1829	.3597	.2321	.1089
		2.	.2194	.1656	.2897	.2304	.1227
			3.	.8085	.4480	.4630	.0688
				4.	.3929	.4771	.0424
					5.	.7400	–.0164
						6.	.0317

b. Girls

N = 206

	1	*2*	*3*	*4*	*5*	*6*	*7*
Sigma ...	107.2950	20.8088	87.6343	11.9300	87.3896	15.3078	1.7342
Mean	179.4612	36.2330	166.4320	14.7039	178.5146	17.4854	6.3301
	1.	.8403	.1054	.1143	.2297	.2487	.0186
		2.	.0010	.1551	.1562	.2753	.0138
			3.	.7254	.4272	.3533	.0900
				4.	.3762	.5193	.1030
					5.	.7717	.0455
						6.	.0550

R.W.B. JACKSON LIBRARY
3 0005 03037626 6

150.8
A673
no. 161
1934
Archives of psychology

150.8
A673
no. 161
1934
Archives of psychology

ARCHIVES OF PSYCHOLOGY

List of numbers, continued from inside front cover

85. Experimental Investigation of Recovery from Work: S. L. Crawley. $1.25.
86. Facilitation and Inhibition: T. N. Jenkins. $1.00.
87. Variability of Performance in the Curve of Work: J. D. Weinland. $1.00.
88. Mental Hygiene Inventory: S. D. House. $1.50.
89. Mental Set and Shift: A. T. Jersild. $1.25.
90. Experimental Investigation of Rest Pauses: C. W. Manzer. $1.25.
91. Routine and Varying Practice as Preparation for Adjustment to a New Situation: L. W. Crafts. $1.00.
93. Speed and Other Factors in "Racial" Difference: O. Klineberg. $1.50.
94. Relation of Reaction Time to Intelligence, Memory, and Learning: V. W. Lemmon. 80c.
95. Is the Latent Time in the Achilles Tendon Reflex a Criterion of Speed in Mental Reactions? G. H. Rounds. $1.25.
96. Predictive Value of Tests of Emotional Stability Applied to College Freshmen: E. G. Flemming. $1.00.
97. Vocabulary Information Test: A. L. Weeks. $1.00.
98. Effect of Temporal Arrangements of Practice on the Mastery of an Animal Maze: S. A. Cook. 80c.
99. Recognition Time as a Measure of Confidence: G. H. Seward. $1.00.
100. Precision and Accuracy: G. W. Hartman. 80c.
101. Group Test of Home Environment: E. M. Burdick. $1.50.
102. Effect of Material on Formal Syllogistic Reasoning: M. C. Wilkins. $1.25.
103. Effect of Incentives on Accuracy of Discrimination: H. C. Hamilton. $1.25.
104. Validity of Norms with Special Reference to Urban and Rural Groups: M. E. Shimberg. $1.25.
105. Blood Pressure Changes in Deception: M. N. Chappell. 80c.
106. Experimental Comparison of Psycho-physical Methods: W. N. Kellogg. $1.25.
107. Measurement of Verbal and Numerical Abilities: M. M. R. Schneck. $1.00.
108. Perseverative Tendency in Pre-School Children. A Study in Personality: H. M. Cushing. $1.00.
109. Preliminary Study of the Effect of Training in Junior High School Shop Courses: L. D. Anderson. 80c.
110. Music Appreciation: M. J. Adler. $1.50.
111. Motivation in Fashion: E. B. Hurlock. $1.00.
112. Equality Judgments in Psychophysics: W. N. Kellogg. $1.00.
113. Illusions in the Perception of Short Time Intervals: N. Israeli. 80c.
114. Further Studies of the Reading-Recitation Process in Learning: Skaggs, Grossman, Krueger & Krueger. 80c.
115. Factors Affecting the Galvanic Reflex: R. C. Davis. $1.00.
116. Infant's Feeding Reactions During the First Six Months: R. Ripin. 80c.
117. Measurement of Mental Deterioration: H. Babcock. $1.25.
118. Phenomenon of Postural Persistence: L. S. Selling. $1.00.
119. American Council on Education Rating Scale: F. F. Bradshaw. $1.00.
120. Group Factor in Immediate Memory: A. Anastasi. $1.00
121. Individual Differences in the Sense of Humor and Temperamental Differences: P. Kambouropoulou. $1.00.
122. Suggestibility in Normal and Hypnotic States: G. W. Williams. $1.00.
123. Analytical Study of the Conditioned Knee-Jerk: G. R. Wendt. $1.25.
124. Race Differences in the Organization of Numerical and Verbal Abilities: J. W. Dunlap. $1.25.
125. Errors of Measurement and Correlation: E. E. Cureton. $1.25.
126. Experience Factors, Test Scores and Efficiency of Women Office Workers: N. Bird. $1.00.
127. Delayed Reactions of Infants: C. N. Allen. 80c.
128. Factors Measured by the Thorndike Intelligence Examination for H. S.: J. G. Peatman. $1.00.
129. Educational Success and Failure in Supernormal Children: J. Regensburg. $1.75.
130. Effect of Practice on Visual Perception of Form: J. P. Seward. $1.00.
131. Relation to College Grades of Some Factors other than Intelligence: D. Harris. 80c.
132. Study of Psychological Differences Between "Racial" and National Groups in Europe: O. Klineberg. $1.00.
133. Emotional Differences of Delinquent and Non-Delinquent Girls of Normal Intelligence: A. Courthial. $1.25.
134. Learning and Retention of Pleasant and Unpleasant Activities: H. Cason. $1.25.
135. Investigation of Brightness Constancy: R. B. MacLeod. $1.25.
136. The Rorschach Test Applied to Feeble-Minded Group: S. J. Beck. $1.00.
137. Retention after Intervals of Sleep and of Waking: E. B. van Ormer. $1.00.
138. Stimulus Temperature and Thermal Sensation: F. Heiser. $1.00.
139. Energy Cost Measurements on Curve of Work: H. J. Schubert. $1.00.
140. Technique for the Measurements of Attitudes: R. Likert. 80c.
141. Speed Factor in Mental Tests: P. H. DuBois. 80c.
142. Further Studies on the Memory Factor: A. Anastasi. $1.00.
143. An Experimental Study on Variability of Learning: S. E. Asch. $1.00.
144. Development of Inventory for Measurement of Inferiority Feelings at H. S. Level: R. B. Smith. $1.50.
145. The Psychological Effects of Oxygen Deprivation: R. A. McFarland. $1.50.
146. Relation of Subliminal to Supraliminal Learning: O. A. Simley. 80c.
147. Effects of Noise upon Certain Psychological Processes: F. L. Harmon. $1.25.
148. Conditioned Responses in Children: G. S. Razran. $1.50.
149. Resemblance of Parents and Children in General Intelligence: M. C. Outhit. $1.00.
150. Influence of Oral Propaganda Material . . . : K-C. Chen. 80c.
151. Negative or Withdrawal Attitude: H. Pallister. 80c.
152. Judgment in Absolute Units as a Psychophysical Method: J. Bressler. $1.00.
153. Visual Illusions in the Chick: C. N. Winslow. $1.25.
154. Measuring Teaching Efficiency Among College Instructors: G. W. Hartmann. 80c.
155. Psychogalvanic Responses in Arithmetical Work: R. Sears. $1.00.
156. Group Factors in Mental Tests: M. Smith. $1.00.
157. . . . Brightness Discrimination Habit in the Chick: A. B. Wood. 80c.
158. Accuracy of perception of visual musical stimuli: A. Roe. $1.00.
159. Effect of Practice Upon Individual Differences: R. E. Perl. $1.00.
160. Antagonistic Muscle Action During the Initiatory Stages of Voluntary Effort: D. J. Wilson. $1.00.
161. Verbal, Numerical and Spatial Abilities of Young Children: B. Schiller. $1.00.

ONTARIO COLLEGE OF EDUCATION
UNIVERSITY OF TORONTO

DIRECTORY OF AMERICAN PSYCHOLOGICAL PERIODICALS

American Journal of Psychology—Ithaca, N. Y.; Cornell University.
Subscription $6.50. 624 pages annually. Edited by M. L. Washburn, K. M. Dallenbach, Madison Bentley, and E. G. Boring. Quarterly. General and experimental psychology. Founded 1887.

Journal of Genetic Psychology—Worcester, Mass.; Clark University Press.
Subscription $14.00 per year; $7.00 per volume. 1000 pages annually. (2 volumes.) Edited by Carl Murchison. Quarterly. Child behavior, animal behavior, and comparative psychology. Founded 1891.

Psychological Review—Princeton, N. J.; Psychological Review Company.
Subscription $5.50. 540 pages annually. Edited by Howard C. Warren. Bi-monthly. General psychology. Founded 1894.

Psychological Monographs—Princeton, N. J.; Psychological Review Company.
Subscription $6.00 per volume. 500 pages. Edited by Herbert S. Langfeld. Published without fixed dates, each issue one or more researches. Founded 1895.

Psychological Index—Princeton, N. J.; Psychological Review Company.
Subscription $4.00. 300–400 pages. Edited by Walter S. Hunter and R. R. Willoughby. An annual bibliography of psychological literature. Founded 1895.

Psychological Bulletin—Princeton, N. J.; Psychological Review Company.
Subscription $6.00. 720 pages annually. Edited by Edward S. Robinson. Monthly (10 numbers). Psychological literature. Founded 1904.

Archives of Psychology—New York, N. Y.; Columbia University.
Subscriptions $6.00. 500 pages per volume. Edited by R. S. Woodworth. Without fixed dates, each number a single experimental study. Founded 1906.

Journal of Abnormal and Social Psychology—Eno Hall, Princeton, N. J.; American Psychological Association.
Subscription $5.00. 448 pages annually. Edited by Henry T. Moore. Quarterly. Abnormal and social. Founded 1906.

Psychological Clinic—Philadelphia, Pa.; Psychological Clinic Press.
Subscription $3.00. 288 pages. Edited by Lightner Witmer. Without fixed dates. (9 numbers). Orthogenics, psychology, hygiene. Founded 1907.

Psychoanalytic Review—Washington, D. C.; 3617 10th St., N. W.
Subscription $6.00. 500 pages annually. Edited by W. A. White and S. E. Jeliffe. Quarterly. Psychoanalysis. Founded 1913.

Journal of Experimental Psychology—Princeton, N. J.
Psychological Review Company. 700 pages annually. Experimetal. Subscription $7.00. Bi-monthly. Edited by S. W. Fernberger. Founded 1916.

Journal of Applied Psychology—Athens, Ohio.
Subscription $5.50. 400 pages annually. Edited by James P. Porter. Bi-monthly. Founded 1917.

Journal of Comparative Psychology—Baltimore, Md.; Williams & Wilkins Company.
Subscription $5.00 per volume of 450 pages. Three volumes every two years. Edited by Knight Dunlap and Robert M. Yerkes. Founded 1921.

Comparative Psychology Monographs—Baltimore, Md.; The Johns Hopkins Press.
Subscription $5.00. 400 pages per volume. Knight Dunlap, Managing Editor. Published without fixed dates, each number a single research. Founded 1922.

Genetic Psychology Monographs—Worcester, Mass.; Clark University Press.
Subscription $14.00 per year; $7.00 per volume. 1000 pages annually. (2 volumes.) Edited by Carl Murchison. Monthly. Each number one complete research. Child behavior, animal behavior, and comparative psychology. Founded 1925.

Psychological Abstracts—Eno Hall, Princeton, N. J.; American Psychological Association.
Subscription $6.00. 700 pages annually. Edited by Walter S. Hunter and R. R. Willoughby. Monthly. Abstracts of psychological literature. Founded 1927.

Journal of General Psychology—Worcester, Mass.; Clark University Press.
Subscription $14.00 per year; $7.00 per volume. 1000 pages annually. (2 volumes.) Edited by Carl Murchison. Quarterly. Experimental, theoretical, clinical, and historical psychology. Founded 1927.

Journal of Social Psychology—Worcester, Mass.; Clark University Press.
Subscription $7.00. 500 pages annually. Edited by John Dewey and Carl Murchison. Quarterly. Political, racial, and differential psychology. Founded 1929.